Quantum Information and Convex Optimization

Von der Fakultät für Elektrotechnik, Informationstechnik, Physik
der Technischen Universität Carolo-Wilhelmina
zu Braunschweig
zur Erlangung des Grades eines
Doktors der Naturwissenschaften
(Dr. rer. nat.)
genehmigte
D i s s e r t a t i o n

von Michael Reimpell
aus Hildesheim

1. Referent: Prof. Dr. Reinhard F. Werner
2. Referent: PD Dr. Michael Keyl
eingereicht am: 22.05.2007
mündliche Prüfung (Disputation) am: 25.09.2007
Druckjahr: 2008

Vorveröffentlichungen der Dissertation

Teilergebnisse aus dieser Arbeit wurden mit Genehmigung der Fakultät für Elektrotechnik, Informationstechnik, Physik, vertreten durch den Mentor der Arbeit, in folgenden Beiträgen vorab veröffentlicht:

Publikationen

[1] M. Reimpell and R. F. Werner. *Iterative Optimization of Quantum Error Correcting Codes.* Physical Review Letters **94** (8), 080501 (2005); selected for[1]: Virtual Journal of Quantum Information **5** (3) (2005); arXiv:quant-ph/0307138.

[2] M. Reimpell, R. F. Werner and K. Audenaert. *Comment on "Optimum Quantum Error Recovery using Semidefinite Programming".* arXiv:quant-ph/0606059 (2006).

[3] M. Reimpell and R. F. Werner. *A Meaner King uses Biased Bases.* arXiv:quant-ph/0612035 (2006).[2]

[4] O. Gühne, M. Reimpell and R. F. Werner. *Estimating Entanglement Measures in Experiments.* Physical Review Letters **98** (11), 110502 (2007); arXiv:quant-ph/0607163.

Tagungsbeiträge

1. M. Reimpell. *Optimal codes beyond Knill & Laflamme?.* Vortrag. Ringberg Meeting, Rottach-Egern, Mai 2003.
2. M. Reimpell and R. F. Werner. *An Iteration to Optimize Quantum Error Correcting Codes.* Vortrag, QIT-EQIS, Kyoto, Japan, September 2003.
3. M. Reimpell. *Quantum error correction as a semi-definite programme.* Vortrag. 26th A2 meeting, Potsdam, November 2003.
4. D. Kretschmann, M. Reimpell and R. F. Werner. *Distributed systems.* Poster.

[1] The Virtual Journal, which is published by the American Physical Society and the American Institute of Physics in cooperation with numerous other societies and publishers, is an edited compilation of links to articles from participating publishers, covering a focused area of frontier research. You can access the Virtual Journal at `http://www.vjquantuminfo.org`.

[2] Erschienen in: Physical Review A **75** (6), 062334 (2007); Virtual Journal of Quantum Information **7** (7) (2007)

Kolloqium des DFG-Schwerpunktprogramms “Quanten-Informationsverarbeitung”, Bad Honnef, Januar 2004.

5. D. Kretschmann, M. Reimpell and R. F. Werner. *Quantum capacity and coding.* Poster. Kolloqium des DFG-Schwerpunktprogramms “Quanten-Informationsverarbeitung”, Bad Honnef, Januar 2004.

6. M. Reimpell and R. F. Werner. *Quantum Error Correcting Codes - Ab Initio.* Vortrag. Frühjahrstagung der Deutschen Physikalischen Gesellschaft, München, März 2004.

7. M. Reimpell. *Introduction to Quantum Error Correction.* Vortrag. A2 - The Next Generation meeting (29th A2 meeting), Potsdam, Juli 2004.

8. M. Reimpell and R. F. Werner. *Iterative Optimization of Quantum Error Correcting Codes.* Poster. Quantum Information Theory: Present Status and Future Directions, Cambridge, August 2004.

9. M. Reimpell and R. F. Werner. *Tough error models.* Vortrag. Frühjahrstagung der Deutschen Physikalischen Gesellschaft, Berlin, März 2005.

10. D. Kretschmann, M. Reimpell, and R. F. Werner. *Quantitative analysis of basic concepts in Quantum Information Theory by optimization of elementary operations.* Poster. Abschlusskolloquium des DFG-Schwerpunktprogramms “Quanten-Informationsverarbeitung”, Bad Honnef, Juni 2005.

11. T. Franz, D. Kretschmann, M. Reimpell, D.Schlingemann and R. F. Werner. *Quantum cellular automata.* Poster. Abschlusskolloquium des DFG-Schwerpunktprogramms “Quanten-Informationsverarbeitung”, Bad Honnef, Juni 2005.

12. M. Reimpell. *Postprocessing tomography data via conic programming.* Vortrag. Arbeitsgruppe Experimental Quantum Physics LMU München, München, Februar 2006.

13. M. Reimpell and R. F. Werner. *Fitting channels to tomography data.* Vortrag. Frühjahrstagung der Deutschen Physikalischen Gesellschaft, Frankfurt, März 2006.

Abstract

This thesis is concerned with convex optimization problems in quantum information theory. It features an iterative algorithm for optimal quantum error correcting codes, a postprocessing method for incomplete tomography data, a method to estimate the amount of entanglement in witness experiments, and it gives necessary and sufficient criteria for the existence of retrodiction strategies for a generalized mean king problem.

keywords: quantum information, convex optimization, quantum error correction, channel power iteration, semidefinite programming, tough error models, quantum tomography, entanglement estimation, witness operators, mean king, retrodiction.

Summary

This thesis investigates several problems in quantum information theory related to convex optimization. Quantum information science is about the use of quantum mechanical systems for information transmission and processing tasks. It explores the role of quantum mechanical effects, such as uncertainty, superposition and entanglement, with respect to information theory. Convexity naturally arises in many places in quantum information theory, as the possible preparations, processes and measurements for quantum systems are convex sets. Convex optimization methods are therefore particularly suitable for the optimization of quantum information tasks. This thesis focuses on optimization problems within quantum error correction, quantum tomography, entanglement estimation and retrodiction.

Quantum error correction is used to avoid, detect and correct noisy interactions of the information carrying quantum systems with the environment. In this thesis, error correction is considered as an optimization problem. The main result is the development of a monotone convergent iterative algorithm, the channel power iteration, which can be used to find optimal coding and decoding operations for arbitrary noise. In contrast to the common approach to quantum error correction, the algorithm does not make any a priori assumptions about the structure of the encoding and decoding operations. In particular, it does not require any error to be corrected completely, and hence can find codes outside the usual quantum error correction setting. More generally, the channel power iteration can be used for the optimization of any linear functional over the set of quantum channels. The thesis also discusses and compares the power iteration to the use of semidefinite programming for the computation of optimal codes. Both techniques are applied to different noise models, where they find improved quantum error correcting codes, and suggest that the Knill-Laflamme form of error correction is optimal in a worst case scenario. Furthermore, the algorithms are used to provide bounds on the correction capabilities for tough error models, and to disprove the optimality of feedback strategies in higher dimensions.

Quantum tomography is used to estimate the parameters of a given quantum state or quantum channel in the laboratory. The tomography yields a sequence of measurement results. As the measurement results are subject to errors, a direct interpre-

tation can suggest negative probabilities. Therefore, parameters are fitted to these tomography data. This fitting is a convex optimization problem. In this thesis, it is shown how to extend the fitting method in the case of incomplete tomography of pure states and unitary gates. The extended method then provides the minimum and maximum fidelity over all fits that are consistent with the tomography data.

A common way to qualitatively detect the presence of entanglement in a quantum state produced in an experiment is via the measurement of so-called witness operators. In this thesis, convexity theory is used to show how a lower bound on a generic entanglement measure can be derived from the measured expectation values of any finite collection of entanglement witnesses. This is shown, in particular, for the entanglement of formation and the geometric measure of entanglement. Thus, with this method, witness measurements are given a quantitative meaning without the need of further experimental data. More generally, the method can be used to calculate a lower bound on any functional on states, if the Legendre transform of that functional is known. It is proven that the bound is optimal under some conditions on the functional.

In the quantum mechanical retrodiction problem, better known as the mean king problem, Alice has to name the outcome of an ideal measurement on a d-dimensional quantum system, made in one of $(d+1)$ orthonormal bases, unknown to Alice at the time of the measurement. Alice has to make this retrodiction on the basis of the classical outcomes of a suitable control measurement including an entangled copy. In this thesis, the common assumption that the bases are mutually unbiased is relaxed. In this extended setting, it is proven that, under mild assumptions on the bases, the existence of a successful retrodiction strategy for Alice is equivalent to the existence of an overall joint probability distribution for $(d+1)$ random variables, whose marginal pair distributions are fixed as the transition probability matrices of the given bases. This provides a connection between the mean king problem and Bell inequalities. The qubit case is completely analyzed, and it is shown how in the finite dimensional case the existence of a retrodiction strategy can be decided numerically via convex optimization.

Contents

Chapter 1

Introduction

Quantum information is an emerging interdisciplinary research field that combines ideas of physics, mathematics, information theory and computer science. It is based on the ability to control quantum systems like photons, atoms, or ions, with the purpose to use them for processing and transmission of information. Quantum information promises or has already offered exciting new possibilities, such as exponential speedup for algorithms, compared to all known classical ones. It allows the generation of cryptographic keys, whose security relies on laws of nature only. Moreover, it offers a chance to by-pass the exponential explosion of the problem size in simulations of quantum systems by using quantum systems themselves as simulators for other quantum systems.

From the point of view of a mathematical physicist, quantum information is the reconsideration of the foundations of quantum mechanics in an information theoretical context. In classical information theory, information is encoded into messages. In the simplest case, these are just strings of bits, which can either be in the state 0 or 1. Classical information theory does not distinguish between different physical carriers of a message, as, in principle, the message can be perfectly transferred and copied between these carriers. However, this is no longer true if the physical carrier is a quantum system. In quantum information theory, the analog of a bit is a qubit. A qubit is a two-level quantum system, for example, the ground and excited state of an atom. Not only can the qubit be in all superpositions of the ground and excited state, the available operations on qubits are fundamentally different from the available operations on bits. Since every measurement of a quantum state introduces some disturbance, and a quantum state cannot be identified in a single measurement, copying of an unknown quantum state is impossible [5]. This fact can also be seen as a consequence of Heisenberg's uncertainty principle. If we could copy an unknown quantum state, we could do a joint measurement of position and momentum by just measuring the position on the original state and the momentum on the clone. Quantum information has unique features such as entanglement, a correlation

between qubits that is stronger than it could be between any classical information carriers. This adds new possibilities in the handling of information theoretical tasks. Like classical information theory, quantum information theory does not differentiate between physical carriers, as long as they are quantum systems. Furthermore, the questions posed in quantum information theory are often similar to the questions in classical information theory. Yet, typically, the answers are different.

Convex optimization is an active subfield of mathematical optimization. It includes least squares, linear programming, and entropy maximization and has a variety of applications in science, engineering, and finance. The concept of convexity is closely related to mixtures and expectation values. Therefore, it plays a major role in quantum mechanics, or, more generally, in any statistical theory. Although there is no analytical formula for the solution of general convex optimization problems, they can reliably and efficiently be solved numerically in many cases. Often, one can even guarantee that the numerical result found is the global optimum within any desired accuracy. Furthermore, every convex optimization problem has an associated dual problem, which sometimes has an interesting interpretation in terms of the original problem. Thus, convex analysis adds another point of view to the problem at hand, and does not merely provide practical algorithms. Even for nonconvex optimization problems, convex optimization methods provide lower bounds on the optimal value by using relaxation techniques. It turns out that optimization problems in quantum information theory can often be formulated as semidefinite programs, a special well-known class of convex optimization problems. These problems are particularly accessible to numerical inspections. However, methods of convex optimization are not yet commonly used in quantum information theory.

Quantum error correction is the umbrella term for techniques to avoid, detect, and correct noise in quantum systems. It is often seen as one of the fundamental technological bases for the scaling of quantum computers. Quantum noise are all processes that corrupt the designated evolution of the quantum system, usually involving nonreversible interactions with the environment. Especially for large systems or long computations, these errors can accumulate and thereby inhibit quantum communication or quantum computation. The usual quantum error correction setting is to protect a subspace of a larger Hilbert space that is subjected to noise. This protected subspace is then used to transfer quantum information. For example, one logical qubit is encoded into several qubits. However, since the state, or equivalently, the quantum data, of such a qubit is unknown, it cannot be copied. So classical codes based on redundancy cannot be applied in the quantum setting. Furthermore, since every measurement to gain information about the current quantum state introduces some perturbation [6], classical codes that condition on the state cannot be applied either. Nevertheless, the classical idea of codewords can be transferred to the quantum case. For example, CSS codes [7, 8] combine two classical linear codes, which meet some technical requirements, to a quantum code.

These codes increase the distinguishability of the codewords, i. e., the possible states of the encoded qubit in the undisturbed case. The detection and correction of errors using linear codes is based on parity checks, so no information about the states themselves is gained. CSS codes are a subclass of so called stabilizer codes [9], the prevalent form of quantum error correcting codes. Stabilizer codes are designed to correct a limited number of specific quantum errors perfectly. Codes that are based on codewords can perfectly correct a given set of quantum errors if and only if they meet the Knill-Laflamme condition [10]. Furthermore, these codes are optimal for asymptotic questions [11]. In total, the theory of perfect error correction with codewords is well established, and already led to experimental realizations [12, 13]. On the other hand, little is known for the case of approximate error correction, although few quantum errors can be corrected perfectly. While perfect error correcting codes can be used to approximately correct these errors in some cases, it is not known in general, whether there are better approximate codes. Furthermore, there is no equivalent for the Knill-Laflamme condition that decides whether noise reduction is possible at all. Also, it is not known if the special form of codeword based encoding is optimal for finite dimensional systems. For some error sources, such as amplitude damping due to the spontaneous emission of a photon, approximate error correcting codes have been found for a fixed choice of dimensions [14]. In this thesis, quantum error correction is regarded as the convex optimization problem to find the best possible code for a given quantum noise. In particular, it will not be assumed that the encoding is based on codewords. Also, for approximate error correction, it will not be required that a code corrects any error perfectly. This corresponds to an ab-initio approach to quantum error correction. Furthermore, bounds on the ability to perfectly correct arbitrary noise are studied in the case that the interaction of the system with its environment is limited.

Quantum tomography describes methods to identify quantum states and quantum channels in the laboratory. For a quantum state produced by a preparation device, several different measurements have to be successively made on the output of that device in order to estimate all parameters of the produced quantum state. These parameters form the density operator of the state, which is a complete characterization of the physical system. For a quantum channel, that is, a device that also takes an input, a combination of preparation for the input and measurements of the output is used to determine the channel parameters. Since quantum channels describe the dynamics of the system, the tomography of quantum channels is also known as process tomography. Due to statistical and systematic errors, a direct interpretation of the obtained measurement results would often result in parameters that do not have a physical meaning, as, for example, they would correspond to negative probabilities. Therefore, parameters are usually assigned from the measurement results using a maximum likelihood estimator [15, 16], which is a convex optimization problem [17, 16, 18]. It is typically assumed that the tomography data

is complete in the sense that, apart form the noise, it specifies all parameters of the given system. However, since the measurements available for a quantum system are limited by the possible interactions with the system, a complete tomography may not be feasible. One goal of this thesis is to extend the convex optimization problem of state and channel tomography to the case of incomplete tomography. Since several parameter sets would reproduce the measured data, this extension should provide the maximum and minimum fidelity of these sets with the designated system, for example, a pure state or unitary time evolution.

Entanglement is a correlation between quantum mechanical systems that is stronger than it could be between classical systems. It is a key resource in quantum information science that is essential to many quantum computational tasks such as teleportation or dense coding. This correlation of quantum systems, for example, two entangled particles, is independent of the spatial separation, and the measurement of a particle at one place can instantly predetermine the outcome of a measurement at the other particle at another place. Due to this combination, Einstein referred to entanglement as "spooky" action at a distance [19]. However, no information transfer is possible based on entanglement alone, so this spooky action cannot be used for superluminal communication. Yet, entanglement excludes the existence of local hidden variables, neither particle can be completely characterized by local parameters. This fact can be described mathematically via Bell inequalities [20] and was verified in a series of experiments by Aspect *et al.* [21, 22, 23]. A common way to test for entanglement in an experiment is via the measurement of so-called witness operators [24], where a negative expectation value witnesses the presence of entanglement. However, so far witness operators are only used for the mere detection of entanglement. As a resource, the amount of entanglement in a given quantum system is of great interest as well. Therefore entanglement measures have been invented (see [25] for a survey). One goal of this thesis is to use convexity theory to give measurements of witness operators a quantitative meaning in terms of entanglement measures.

Retrodiction in quantum information theory is the problem to reconstruct the values of measurements which can no longer be obtained from the system itself. It is often told as the mean king tale [26, 27, 28, 29]. Alice has to ascertain the result of a basis measurement made by the king. The king randomly chooses this basis from a set of, in general, non-commuting bases. Alice is allowed to do the initial preparation and to make a final measurement on the system. Although she knows all possible bases, Alice gains information about the king's choice only after she has no longer access to the system. Retrodiction is interesting with respect to quantum cryptography, where such random choices are frequently made. Indeed, it is known that Alice can retrodict the king's measurement result with certainty, when the bases are mutually unbiased [30, 26, 27, 31, 28, 29]. Mutually unbiased means that the measurement of one basis does not give any information about the

outcome of a measurement of another basis, which fallaciously suggests that this is a particular mean choice. In this thesis, the assumption that the bases are mutually unbiased is relaxed. In this more general setting it is studied whether Alice can find a successful retrodiction strategy, and the mean king problem is shown to be a convex optimization problem.

Each of the above goals is closely related to convex optimization. The basic concepts of quantum information and convex optimization are briefly presented in Chapter 2. Chapter 4 about postprocessing tomography data depends on definitions from Chapter 3 about quantum error correction. The other chapters are rather self-contained. As the chapters address different topics in quantum information theory, each comes with its own conclusion.

Chapter 2

Basic Concepts

This chapter gives a brief survey of the mathematical framework of quantum information theory and convex optimization. The survey is restricted to those topics that are required for the understanding of the later chapters. A good nontechnical introduction to quantum information is given by Werner [32]. For a short review of the theory and applications of semidefinite programming, the most frequently used convex optimization method in this thesis, I recommend the text by Vandenberghe and Boyd [33]. A more complete treatment of quantum information can be found in [34, 35]. The mathematical framework of operator algebras is presented in the textbooks [36, 37, 38]. Convex optimization is addressed in more detail in [39, 40, 41].

2.1 Proper Cones

Before we introduce quantum information theory, we have a quick glance at the mathematical concepts of convexity and cones. Convexity of a set means that the line segment between any two elements of the set lies in the set as well.

2.1.1 Definition (Convex Set)**.** A set C with elements of a vector space is convex, if we have

$$\lambda x + (1 - \lambda)y \in C$$

for any $x, y \in C$ and any $0 \leq \lambda \leq 1$.

It immediately follows that if C is convex, $x_i \in C$ for $i = 1, \ldots, n$ and $0 \leq \lambda_i \leq 1$ with $\sum_{i=1}^{n} \lambda_i = 1$, then we have $\sum_{i=1}^{n} \lambda_i x_i \in C$. Such a sum is called a convex combination with weights λ_i.

A convex combination can be interpreted as a mixture or weighted average, and therefore has a natural connection to probability distributions.

If a convex set is closed and bounded, then it can be generated via all convex combinations of its extreme points. A point x of a convex set is called an extreme point if and only if x cannot be expressed as a convex combination $x = \lambda y + (1-\lambda)z$, $0 < \lambda < 1$, except by taking $y = z = x$. Extreme points can be thought of as corners of the set.

The definition of convexity can be transferred to functions. A real-valued function f on a vector space is said to be convex if and only if the epigraph

$$\{(x,t)\,|x \in \operatorname{dom} f, f(x) \leq t\},$$

that is, the graph above the function, is a convex set. This is equivalent to the condition that $\operatorname{dom} f$ is a convex set and that

$$f(\lambda x + (1-\lambda)y) \leq \lambda f(x) + (1-\lambda)f(y)$$

for all $x, y \in \operatorname{dom} f$ and all λ with $0 \leq \lambda \leq 1$.

We are also interested in sets, where we can scale the elements without leaving the set.

2.1.2 Definition (Cone)**.** A set C is called a cone, if for all $x \in C$ and all $\lambda \geq 0$ we have $\lambda x \in C$.

If a set C is convex and a cone, we say that C is a convex cone. Combining both properties, we have $\mu\lambda x + \mu(1-\lambda)y \in C$ for all $\mu \geq 0$, $0 \leq \lambda \leq 1$, and all $x, y \in C$. This means that C is a convex cone, if for all $\mu_1, \mu_2 \geq 0$ and all $x, y \in C$ we have

$$\mu_1 x + \mu_2 y \in C.$$

Thus, for any two vectors x and y, the pie slice with apex 0 and edges passing through x and y is in the convex cone as well.

We are mostly interested in so-called proper cones.

2.1.3 Definition (Proper Cone)**.** A convex cone C in a normed vector space is a proper cone, if the following conditions hold.

1. C is closed. That is, the limit of any sequence of vectors in C is in C.
2. C has nonempty interior. That is, there exists a vector such that a ball with strictly positive radius centered at that vector is contained in C.
3. C is pointed. That is, if $x \in C$ and $-x \in C$ then $x = 0$.

Such a proper cone induces a partial ordering "$\geq_C$".

2.1.4 Proposition. *Let a subset C of a vector space V be a proper cone. Then, a partial ordering on V is given by*

$$x \geq_C y \quad \Leftrightarrow \quad x - y \in C.$$

We also say $x >_C y$ if $x - y$ is in the interior of C. We will usually drop the index C, however, the partial ordering should not be confused with an ordinary inequality. For example, there can be elements x and y in the vector space, such that neither $x \geq y$ nor $y \geq x$ holds.

We will only consider cones in vector spaces over a field with a partial ordering, and where the vector space has a scalar product $\langle \cdot | \cdot \rangle$, for example, $C \subset \mathbb{R}^n$. If C is a cone in such a space, then we define the dual cone as

$$C^* = \{ |y\rangle \, | \langle y|x\rangle \geq 0 \text{ for all } |x\rangle \in C \} .$$

Note that the dual cone C^* is always convex, even if the primal cone C is not. Furthermore, if the primal cone C is a proper cone, then the dual cone C^* is also a proper cone and we have $(C^*)^* = C$. Important proper cones are selfdual, meaning that even $C^* = C$. Among them are the nonnegative orthant $\mathbb{R}^n_+$, the positive semidefinite cone S, and the quadratic cone Q. The positive semidefinite cone S is defined as the set of positive semidefinite matrices,

$$S = \left\{ M \in \mathbb{C}^{d \times d} \,\middle|\, M = M^*, \langle v|Mv\rangle \geq 0 \text{ for all } |v\rangle \in \mathbb{C}^d \right\} .$$

The quadratic cone Q is defined as

$$Q = \left\{ (x, y) \in \mathbb{R} \times \mathbb{C}^{d-1} \, | x \geq \|y\| \right\} .$$

In particular, the semidefinite cones are frequently used in quantum information theory, where we will use intersections of hyperplanes with these cones to model convex constraints.

2.2 Quantum Information Theory

Quantum information theory, just like classical information theory, is a statistical theory. Thus, in order to test its predictions, experiments have to be often repeated and the predicted probabilities have to be compared to the relative frequencies of the outcomes. In quantum information theory, a statistical experiment is assembled from two components, the preparation procedure and the observation procedure. During the preparation, a physical system is engineered in a distinguished *state*. During the observation, properties of such a state are measured. We will build up all observations from boolean valued measurements and call such a measurement of a truth value an *effect*.

Given a physical system, the mathematical description of such a statistical experiment therefore consists of the following entities: The set $\mathcal{S}$ of all states that a given system can be prepared in, the set $\mathcal{E}$ of all effects that can be measured on the system, and a map

$$\mathcal{S} \times \mathcal{E} \ni (\rho, A) \mapsto \rho(A) \in [0, 1], \tag{2.1}$$

that maps all combinations of states and effects to the probability to obtain the result "true" in the corresponding experiment.

This general scheme does not only apply to quantum systems, but also to classical systems and hybrid systems, that is, systems that have quantum and classical parts. Each of these systems can be characterized by a so-called observable algebra $\mathcal{A}$, where $\mathcal{A}$ is a C*-algebra with identity. A C*-algebra is a complex vector space with associative and distributive multiplication, an involution $\cdot^*$ and a norm $\|\cdot\|$.

The involution, also known as adjoint operation, is an antilinear operation with $(AB)^* = B^*A^*$ and $(A^*)^* = A$ for all $A, B \in \mathcal{A}$. Antilinear means that $(A+B)^* = A^* + B^*$ and $(\alpha A)^* = \overline{\alpha} A^*$, where $\overline{\alpha}$ denotes the complex conjugate of the scalar $\alpha \in \mathbb{C}$.

The norm satisfies $\|AB\| \leq \|A\|\|B\|$ and $\|A^*A\| = \|A\|^2$ (and thus $\|A^*\| = \|A\|$) for all $A, B \in \mathcal{A}$, in addition to the positive definiteness ($\|A\| \geq 0$ and $\|A\| = 0$ if and only if $A = 0$), the positive homogeneity ($\|\alpha A\| = |\alpha|\|A\|$ for all scalars α), and the triangle inequality ($\|A + B\| \leq \|A\| + \|B\|$). Furthermore $\mathcal{A}$ is closed under this norm. We will use $\mathbb{1}$ as symbol for the identity of $\mathcal{A}$.

For the description of quantum mechanical systems we have $\mathcal{A} = \mathcal{B}(\mathcal{H})$, the algebra of bounded operators on a Hilbert space $\mathcal{H}$. In this thesis, only finite dimensional Hilbert spaces are considered if not stated otherwise. Thus, we can take $\mathcal{H} = \mathbb{C}^d$, so $\mathcal{B}(\mathcal{H})$ corresponds to the matrix algebra of complex $d \times d$ matrices.

A special subset of elements of $\mathcal{A}$ we will frequently refer to, is the set of positive operators.

2.2.1 Definition (Positive Operators)**.** Let $\mathcal{A}$ be a C*-algebra. An operator $A \in \mathcal{A}$ is called *positive*, if A is selfadjoint, that is, $A = A^*$, and the spectrum of A is a subset of the positive half-line $\mathbb{R}_+$.

In $\mathcal{B}(\mathcal{H})$, we have the following equivalent characterizations of positivity (see Thm. 2.2.12 [36]).

2.2.2 Proposition. *Let $A \in \mathcal{B}(\mathcal{H})$, where $\mathcal{H}$ is a finite dimensional Hilbert space. Then the following conditions are equivalent.*

1. *A is positive.*
2. *$\langle\psi|A\psi\rangle \geq 0$ for all $|\psi\rangle \in \mathcal{H}$.*

3. $A = B^*B$ *for a (unique) selfadjoint operator* $B \in \mathcal{B}(\mathcal{H})$.

The set of positive operators is a proper cone (Prop. 2.2.11 [36]) and therefore induces a partial ordering.

2.2.3 Proposition (Positive Cone)**.** *Let* $A, B, C \in \mathcal{A}$. *We say* $A \geq B$, *if* $A - B$ *is positive. The relation "*$\geq$*" defines a partial ordering, i. e., it is reflexive, antisymmetric and transitive. In particular, if* $A \geq 0$ *and* $A \leq 0$, *then we have* $A = 0$. *Furthermore, the following implications are valid (see Prop. 2.2.13 [36]):*

1. *if* $A \geq B \geq 0$ *then* $\|A\| \geq \|B\|$;
2. *if* $A \geq B \geq 0$ *then* $C^*AC \geq C^*BC \geq 0$ *for all* C;

Every operator in $\mathcal{A}$ can be expressed as a linear combination of positive operators.

2.2.4 Proposition. *Let* $\mathcal{A}$ *be a* C^**-algebra with identity. Then, every element* $A \in \mathcal{A}$ *has a decomposition of the form*

$$A = +A_1 - A_2 + iA_3 - iA_4,$$

with positive elements A_i.

Proof. We can write A as the linear combination $A = A_r + iA_i$ of the selfadjoint operators $A_r = 1/2(A+A^*)$ and $A_i = 1/(2i)(A-A^*)$. Furthermore, every selfadjoint element of $\mathcal{A}$ can be decomposed into a linear combination of positive operators (Prop. 2.2.11 [36]). So for $B^* = B \in \mathcal{A}$ we have[1] $B = B_+ - B_-$ with $B_\pm = 1/2(|B| \pm B) \geq 0$. Combining the two decompositions leads to the linear decomposition of an arbitrary operator of $\mathcal{A}$ into positive operators. ■

With the notion of positivity, we can now define the set of states $\mathcal{S}$ and effects $\mathcal{E}$ for the quantum case $\mathcal{A} = \mathcal{B}(\mathcal{H})$. The set of effects is defined as

$$\mathcal{E} = \{A \in \mathcal{B}(\mathcal{H}) \,|\mathbb{1} \geq A \geq 0\}.$$

The set of states is defined using the dual space,

$$\mathcal{S} = \{\varrho \in \mathcal{B}(\mathcal{H})^* \,|\varrho \geq 0, \varrho(\mathbb{1}) = 1\}.$$

Given the state ϱ, the probability to measure the effect A is $\varrho(A)$. Observe that the combination of the constraints of both sets ensure that the result of an experiment is always in the interval $[0, 1]$, and hence can be interpreted as a probability. Both sets are convex and are completely characterized by their extreme points. As we restrict the dimension of the Hilbert space $\mathcal{H}$ to be finite dimensional, we can apply Riesz

[1] The operator $|B|$ is defined below.

Theorem [42] and identify the state ϱ with an operator ρ using the Hilbert-Schmidt scalar product $\langle A|B\rangle = \mathrm{tr}(A^*B)$,

$$\varrho(A) = \langle \rho|A\rangle = \mathrm{tr}(\rho A).$$

With this identification, the set of states becomes $\{\rho \in \mathcal{B}(\mathcal{H}) \,|\rho \geq 0, \mathrm{tr}(\rho) = 1\}$. The extreme points characterizing this set are the one-dimensional projections $|\psi\rangle\langle\psi|$, which are called pure states. Therefore, we will also refer to vectors $|\psi\rangle$ as pure states. Note that the trace norm $\|\rho\|_1 = \mathrm{tr}\,|\rho|$ is used for states, while the operator norm $\|A\| = \sup_{\psi\in\mathcal{H},\|\psi\|=1} \|A\psi\|$ is used for effects. Here $|\rho|$ is defined as $|\rho| = \sqrt{A^*A}$ in terms of the functional calculus:

2.2.5 Definition. Let f be a complex function, $f\colon \mathbb{C} \to \mathbb{C}$, and $A \in B(\mathbb{C}^d)$ be an operator with eigenvalue decomposition $A = \sum_{i=1}^d \lambda_i |e_i\rangle\langle e_i|$. Then, $f(A)$ is defined as

$$f(A) = \sum_{i=1}^{d} f(\lambda_i)|e_i\rangle\langle e_i|.$$

Since A^*A is positive, an eigenvalue decomposition exists, and thus, $\sqrt{A^*A}$ defines $|\rho|$ for all operators $A \in B(\mathbb{C}^n)$.

By associating an effect with every possible outcome of a measurement, we can describe more general measurements.

2.2.6 Definition (POVM)**.** A positive operator valued measure (POVM) for a finite set X of measurement outcomes is a set of effects E_x, $x \in X$, such that $\sum_{x\in X} E_x = \mathbb{1}$.

Often, we have that E_x are projections. In this case we call the set of effects a projection valued measure (PVM). For example, the eigenvalue decomposition $A = \sum_i \lambda_i |e_i\rangle\langle e_i|$ corresponds to the PVM with operators $|e_i\rangle\langle e_i|$. Moreover, we can interpret any hermitian operator $A \in B(\mathcal{H})$ as an observable with the expectation value $\mathrm{tr}(\rho A) = \sum_i \lambda_i \,\mathrm{tr}(\rho|e_i\rangle\langle e_i|) = \sum_i \lambda_i \langle e_i|\rho|e_i\rangle$ for a given state ρ.

We can compose systems out of different subsystems via the tensor product. Let $\mathcal{H}$ and $\mathcal{K}$ be finite dimensional Hilbert spaces that correspond to two subsystems with observable algebras $\mathcal{B}(\mathcal{H})$ and $\mathcal{B}(\mathcal{K})$. Then, the observable algebra for the composite system is given by $\mathcal{B}(\mathcal{H}) \otimes \mathcal{B}(\mathcal{K}) \simeq \mathcal{B}(\mathcal{H} \otimes \mathcal{K})$.

Here $\mathcal{H} \otimes \mathcal{K}$ is defined as the linear span of tensor products $|\psi\rangle \otimes |\phi\rangle$ of vectors $|\psi\rangle \in \mathcal{H}$ and $|\phi\rangle \in \mathcal{K}$. That is, given two bases $\{|\psi_i\rangle\} \subset \mathcal{H}$ and $\{|\phi_j\rangle\} \subset \mathcal{K}$, every vector $|\varphi\rangle \in \mathcal{H} \otimes \mathcal{K}$ can be decomposed as

$$|\varphi\rangle = \sum_{ij} \alpha_{ij} |\psi_i\rangle \otimes |\phi_j\rangle$$

with scalars α_{ij}. The tensor product is a bilinear form, which means that

$$\begin{aligned}\alpha(|\psi\rangle \otimes |\phi\rangle) &= (\alpha|\psi\rangle) \otimes |\phi\rangle = |\psi\rangle \otimes (\alpha|\phi\rangle),\\ (|\psi_1\rangle + |\psi_2\rangle) \otimes |\phi\rangle &= |\psi_1\rangle \otimes |\phi\rangle + |\psi_2\rangle \otimes |\phi\rangle,\\ |\psi\rangle \otimes (|\phi_1\rangle + |\phi_2\rangle) &= |\psi\rangle \otimes |\phi_1\rangle + |\psi\rangle \otimes |\phi_2\rangle,\end{aligned}$$

for scalar α and vectors $|\psi\rangle, |\psi_1\rangle, |\psi_2\rangle \in \mathcal{H}$ and $|\phi\rangle, |\phi_1\rangle, |\phi_2\rangle \in \mathcal{K}$. The scalar product of $\mathcal{H} \otimes \mathcal{K}$ is defined by

$$\langle \psi_1 \otimes \phi_1 | \psi_2 \otimes \phi_2 \rangle = \langle \psi_1 | \psi_2 \rangle_{\mathcal{H}} \langle \phi_1 | \phi_2 \rangle_{\mathcal{K}}$$

Likewise, $\mathcal{B}(\mathcal{H}) \otimes \mathcal{B}(\mathcal{K})$ is the linear span of operators of the form $(A \otimes B)$, with the additional operations

$$\begin{aligned}(A \otimes B)(C \otimes D) &= (AC) \otimes (BD),\\ (A \otimes B)^* &= A^* \otimes B^*.\end{aligned}$$

The operator $A \otimes B$ acts on the vector $|\psi\rangle \otimes |\phi\rangle$ as $A|\psi\rangle \otimes B|\phi\rangle$, which is extended by linearity for general $C \in \mathcal{B}(\mathcal{H}) \otimes \mathcal{B}(\mathcal{K})$ and $|\varphi\rangle \in \mathcal{H} \otimes \mathcal{K}$.

The effect $A \otimes B \in B(\mathcal{H} \otimes \mathcal{K})$ corresponds to the joint measurement of A on the first and B on the second system. The system can be restricted to one of the subsystems using the partial trace.

2.2.7 Definition (Partial Trace)**.** Let $\rho \in \mathcal{B}(\mathcal{H} \otimes \mathcal{K})$ be the state of the composite system, then the partial trace $\mathrm{tr}_{\mathcal{K}}$ that restricts the state to the first subsystem $\mathcal{B}(\mathcal{H})$ is defined by the equation

$$\mathrm{tr}\left(\mathrm{tr}_{\mathcal{K}}(\rho)A\right) = \mathrm{tr}(\rho(A \otimes \mathbb{1}))$$

for all operators $A \in \mathcal{B}(\mathcal{H})$.

Note that $\mathrm{tr}_{\mathcal{K}}(\rho) \in B(\mathcal{H})$ is an operator, and that $\mathbb{1} \in \mathcal{B}(\mathcal{K})$ corresponds to the effect that ignores the output of the measurement on that system.

We can now look at the possible correlations between subsystems. A state $\rho \in \mathcal{B}(\mathcal{H} \otimes \mathcal{K})$ is called correlated, if there are effects $A \in \mathcal{B}(\mathcal{H})$, $B \in \mathcal{B}(\mathcal{K})$, such that

$$\mathrm{tr}\left(\rho(A \otimes B)\right) \neq \mathrm{tr}\left(\mathrm{tr}_{\mathcal{K}}(\rho)A\right) \mathrm{tr}\left(\mathrm{tr}_{\mathcal{H}}(\rho)B\right).$$

This implies that states that are not correlated can be written as a tensor product, $\rho = \rho_1 \otimes \rho_2$. Such states can be prepared using two preparation devices. One locally prepares ρ_1 on the first subsystem, the other locally prepares ρ_2 on the other subsystem. Due to the convexity of the state space, we know that we can also prepare mixtures of such states $\rho = \sum_i \lambda_i \rho_{1,i} \otimes \rho_{2,i}$. This can also be done with local preparation devices by sharing the result of a random number generator between them,

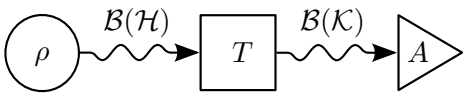

Figure 2.1: Quantum experiment with channel T, preparation ρ and observable A. In the Heisenberg picture, T maps the observable $A \in B(\mathcal{K})$ to the observable $T(A) \in B(\mathcal{H})$. The corresponding channel T_* in the Schrödinger picture maps the state ρ to the state $T_*(\rho) \in B(\mathcal{K})$.

where the random number generator produces output i with probability λ_i. Here, the sharing of the random number leads to correlations between the subsystems. However, not all states of a composite quantum system can be prepared in this way, which brings us to the definition of entanglement.

2.2.8 Definition (Entanglement)**.** A state $\rho \in \mathcal{B}(\mathcal{H} \otimes \mathcal{K})$ is called separable or classical correlated, if it can be written as a convex combination of the form

$$\rho = \sum_i \lambda_i \rho_{1,i} \otimes \rho_{2,i},$$

with weights λ_i and states $\rho_{1,i} \in \mathcal{B}(\mathcal{H})$, $\rho_{2,i} \in \mathcal{B}(\mathcal{K})$. Otherwise, ρ is called entangled.

Entanglement is unique to quantum systems. There is neither entanglement between classical systems, nor is there entanglement between the quantum and classical parts of a hybrid system.

Operations on quantum systems, for example, the free time evolution, are used to process quantum information. As in classical information theory, we describe such a processing step with a channel.

2.2.9 Definition (Quantum Channel)**.** A quantum channel is a linear, completely positive, unital map with

$$T\colon \mathcal{B}(\mathcal{K}) \to \mathcal{B}(\mathcal{H})$$

A map is called *positive*, if it maps positive operators to positive operators. It is called *completely positive*, if this is the case even when the map is only applied to a subsystem. That is, T is completely positive if

$$T \otimes \mathrm{id}_n\colon \mathcal{B}(\mathcal{K}) \otimes \mathcal{B}(\mathbb{C}^n) \to \mathcal{B}(\mathcal{H}) \otimes \mathcal{B}(\mathbb{C}^n)$$

is positive for all $n \in \mathbb{N}$, where id_n denotes the identity map on $\mathcal{B}(\mathbb{C}^n)$. Note that positivity of a channel does not imply that the channel is completely positive. Complete positivity allows to use channels to describe operations local to a subsystem.

A map $T\colon \mathcal{B}(\mathcal{K}) \otimes \mathcal{B}(\mathcal{H})$ is called *unital*, if it maps the unity of $\mathcal{B}(\mathcal{K})$ to the unity of $\mathcal{B}(\mathcal{H})$,

$$T(\mathbb{1}_{\mathcal{K}}) = \mathbb{1}_{\mathcal{H}}.$$

Combined with linearity and complete positivity this ensures that a channel maps effects to effects, possibly on a different quantum system. We will also consider *subchannels*, where we relax the unital condition to $T(\mathbb{1}) \leq \mathbb{1}$. This means that the channel is allowed to sometimes produce no output at all.

A schematic diagram of a quantum information experiment with a channel $T\colon \mathcal{B}(\mathcal{K}) \to \mathcal{B}(\mathcal{H})$ is shown in Figure 2.1. If ρ is the state of the system in $\mathcal{B}(\mathcal{H})$ and A is the effect measured on the system $\mathcal{B}(\mathcal{K})$, then the probability of obtaining the result "true" is

$$\operatorname{tr}\big(\rho\, T(A)\big).$$

We can also describe the processing step using the predual of T. The predual of T is the map $T_*\colon \mathcal{B}(\mathcal{H}) \to \mathcal{B}(\mathcal{K})$, for which

$$\operatorname{tr}\big(T_*(\rho)A\big) = \operatorname{tr}\big(\rho\, T(A)\big).$$

The predual is linear and completely positive. The unital condition of T translates to the condition that T_* is trace preserving,

$$\operatorname{tr}\big(T_*(\rho)\big) = \operatorname{tr}\big(T_*(\rho)\mathbb{1}\big) = \operatorname{tr}\big(\rho\, T(\mathbb{1})\big) = \operatorname{tr}(\rho\mathbb{1}) = \operatorname{tr}(\rho).$$

We will call T the channel in the Heisenberg picture, and refer to T_* as the channel in the Schrödinger picture. Observe that the set of quantum channels between two quantum systems is convex.

Every channel can be represented as follows [43].

2.2.10 Theorem (Stinespring Dilation Theorem)**.** *Let $\mathcal{H}$ and $\mathcal{K}$ be two Hilbert spaces with dimensions* $\dim \mathcal{H}$ *and* $\dim \mathcal{K}$*. Then, every channel $T\colon \mathcal{B}(\mathcal{K}) \to \mathcal{B}(\mathcal{H})$ can be written in the form*

$$T(A) = v^*(A \otimes \mathbb{1}_{\mathcal{D}})v.$$

Here v is an isometry, that is, $v^*v = \mathbb{1}$, and $\mathcal{D}$ is an additional Hilbert space called dilation space*. This form is called the* Stinespring representation *of the map T. For the minimal dilation dimension we have* $\dim \mathcal{D} \leq (\dim \mathcal{H})(\dim \mathcal{K})$*. The minimal Stinespring representation is unique up to unitary equivalence.*

A closely related representation is the Kraus decomposition [44], which is also known as operator-sum representation.

2.2.11 Corollary (Kraus Decomposition)**.** *Let $\mathcal{H}$ and $\mathcal{K}$ be two Hilbert spaces with dimensions* $\dim \mathcal{H}$ *and* $\dim \mathcal{K}$*. Then, every channel $T\colon \mathcal{B}(\mathcal{K}) \to \mathcal{B}(\mathcal{H})$ can be written in the form*

$$T(A) = \sum_{i=1}^{N} t_i^* A\, t_i,$$

with so-called Kraus operators $t_i\colon \mathcal{H} \to \mathcal{K}$. *The number of independent Kraus operators is fixed with* $N \leq (\dim \mathcal{K})(\dim \mathcal{H})$.

Proof. Let $T(A) = v^*(A \otimes \mathbb{1}_{\mathcal{D}})v$ be a Stinespring dilation of the channel T with dilation space $\mathcal{D}$. Consider a family $|\chi_i\rangle\langle\chi_i|$ of one-dimensional projectors with $\sum_i |\chi_i\rangle\langle\chi_i| = \mathbb{1}_{\mathcal{D}}$. A Kraus representation is then given by the operators t_i defined by $\langle\phi|t_i\psi\rangle = \langle\phi \otimes \chi_i|v\psi\rangle$. On the other hand, given a Kraus representation of the channel one can choose an orthonormal basis $|\chi_i\rangle$ to get a Stinespring isometry. The isometry property is implied by the unital condition $T(\mathbb{1}) = \sum_i t_i^* t_i = \mathbb{1}$. As the minimal dilation dimension is fixed, the minimal number of Kraus operators is also fixed. If the Kraus operators aren't linearly independent, one could choose a smaller dilation space and hence a new Kraus decomposition with a smaller number of Kraus operators. Therefore, in Kraus decompositions with minimal number of operators, these operators are linearly independent. As the maximal size of the matrix of such a Kraus operator is $(\dim \mathcal{K})(\dim \mathcal{H})$, the number of linearly independent Kraus operators is below this value. ■

Note that if $T(A) = \sum_i t_i^* A\, t_i$ is a channel in the Heisenberg picture, then the corresponding channel in the Schrödinger picture is $T_*(\rho) = \sum_i t_i \rho\, t_i^*$, since

$$\operatorname{tr}\big(\rho\, T(A)\big) = \sum_i \operatorname{tr}(\rho\, t_i^* A t_i) = \sum_i \operatorname{tr}(t_i \rho\, t_i^* A) = \operatorname{tr}\big(T_*(\rho)\, A\big).$$

Another channel representation that is closely related to the Stinespring Dilation Theorem is the ancilla form, also known as unitary representation.

2.2.12 Corollary (Ancilla Form)**.** *Let* $T_*\colon \mathcal{B}(\mathcal{H}) \to \mathcal{B}(\mathcal{H})$ *be a channel in the Schrödinger picture. Then,* T_* *can be written in the form*

$$T_*(\rho) = \operatorname{tr}_{\mathcal{K}}\big(U(\rho \otimes \rho_{\mathcal{K}})U^*\big),$$

with an additional Hilbertspace $\mathcal{K}$*, a state* $\rho_{\mathcal{K}} \in \mathcal{B}(\mathcal{K})$ *and a unitary operator* U*, that is,* $U^*U = UU^* = \mathbb{1}$.

The proof (e. g., see [34]) is based on the fact that if $T(A) = v^*(A \otimes \mathbb{1}_{\mathcal{K}})v$ is a Stinespring representation of T, then we can extend the isometry $v\colon \mathcal{H} \to \mathcal{H} \otimes \mathcal{K}$ to a unitary operator $U\colon \mathcal{H} \otimes \mathcal{K} \to \mathcal{H} \otimes \mathcal{K}$. Since the time evolution of a closed quantum system is reversible and therefore represented by a unitary operator, we can interpret the ancilla form as the common evolution of the system with the environment $\mathcal{K}$, where the initial state of the composite system is $\rho \otimes \rho_{\mathcal{K}}$. For example, if H is a Hamilton operator that describes the evolution of the composite system, we can choose $U = e^{-iHt/\hbar}$. However, bear in mind that the ancilla form is not unique. A channel can be represented by several different unitary operators U.

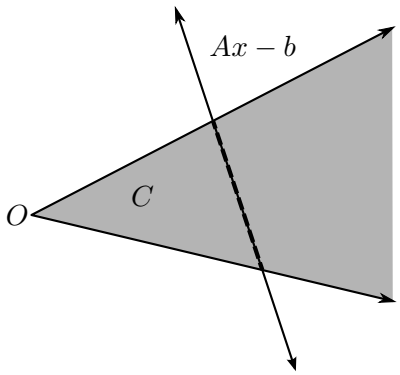

Figure 2.2: Constraint of a conic program. A conic program is the minimization of a linear objective (not depicted) over the intersection (dashed line) of a convex cone C and an affine plane $\{y \,|y = Ax - b\}$. The apex of the cone is the origin O of the underlying vector space.

For a closed quantum system $\mathcal{B}(\mathcal{H})$ with time evolution $U = e^{-iHt/\hbar}$ generated by the Hamilton operator H, we immediately see from the ancilla form that

$$T(|\psi\rangle\langle\psi|) = U|\psi\rangle\langle\psi|U^* = |U\psi\rangle\langle U\psi|.$$

Hence, in this case, T maps pure states to pure states. Moreover, since the set of states is completely characterized by the extreme points, we only have to consider the action of U on vectors $|\psi\rangle \in \mathcal{H}$.

2.3 Convex Optimization

Convex optimization problems are optimization problems, where the objective and the constraints are convex. We will consider optimization problems with linear objective, where the convex constraints can be modeled as intersection of an affine plane with a proper cone. Such a convex optimization problem is called a conic program[2]. In its primal form, it can be written as

$$\min_x \{\langle c|x\rangle \,|Ax - b \geq_C 0\}\,.$$

Here $\{y \,|y = Ax - b\}$ is an affine plane and $\geq_C$ is the partial ordering induced by the cone C. Thus $Ax - b \geq 0$ is the intersection of both. When the scalar product is complex, only the real part is optimized. Figure 2.2 shows an example of the constraint of a conic program.

[2]The term "program" refers to a mathematical optimization problem.

Every primal problem has an associate dual problem. To see this, we look at the constraints in the dual cone. For $\lambda \in C^*$ we have $\langle \lambda | Ax - b \rangle \geq 0$ and thus

$$\langle A^* \lambda | x \rangle \geq \langle \lambda | b \rangle.$$

So for every $\lambda \in C^*$ with $A^* \lambda = c$ leads to a lower bound on the primal objective,

$$\langle A^* \lambda | x \rangle = \langle c | x \rangle \geq \langle \lambda | b \rangle.$$

The dual program is the optimization to find the maximal lower bound,

$$\max_{\lambda} \left\{ \langle \lambda | b \rangle \, | \lambda \geq_{C^*} 0, A^* \lambda = c \right\}.$$

The duality is symmetric, that is, the dual problem is also a conic program, and the dual of the dual problem is equivalent to the primal problem.

An element x is called feasible if it satisfies the constraints, that is, $Ax - b \geq 0$. Likewise, a dual feasible element λ satisfies the dual constraints. For any pair of feasible elements (x, λ), the difference between the two objectives, $\langle c | x \rangle - \langle \lambda | b \rangle$, is called the duality gap. It follows from the above that it is always nonnegative,

$$\langle c | x \rangle - \langle \lambda | b \rangle \geq 0.$$

With appropriate constraint qualification (see Thm. 2.4.1 [40]), we even have strong duality, which means that the duality gap is zero for an optimal pair of primal and dual feasible points. For example, this is the case if the primal or dual problem is bounded and strictly feasible. Strictly feasible means that there exists an element in the interior of the cone that satisfies the constraints. For the primal problems, this means there exists an x such that $Ax - b > 0$. For the dual problem, this means that there exists a $\lambda > 0$ with $A^* \lambda = c$.

So whenever we can find a pair of feasible points (x, λ) with duality gap ε, we know that the objectives $\langle c | x \rangle$ and $\langle \lambda | b \rangle$ lie in an ε-interval around the true global optimum. Thus, the dual point λ certifies the optimality of the primal point x up to ε. This is the main virtue of conic programming, since there exist numerical algorithms that solve both problems for arbitrary small ε. Note that when strong duality holds, we necessarily have $\langle \lambda | Ax - b \rangle = 0$ for any optimal pair (x, λ). This condition is known as complementary slackness.

The duality gap has another interesting implication. If both optimization problems have a solution, the duality gap implies that both optimization problems are bounded. So whenever the primal or dual problem is unbounded and we can find a feasible point for any of the problems, the other problem cannot have a feasible point as well. This is known as Theorem of Alternatives or as Farkas' Lemma.

There are some commonly used conversion techniques for the handling of conic programs. Observe that partial ordering constraints can be translated into equality

constraints and vice versa. A partial ordering constraint $a \geq b$ can be written as equality constraint $a + s = b$ using a slack variable s that is constrained to the cone, that is, $s \geq 0$. On the other hand, an equality constraint $a = b$ can be expressed via the two partial ordering constraints $a \geq b$ and $-a \geq -b$. Due to this fact, there are several different but equivalent formulations of conic programs in the literature.

Another common technique is to introduce an auxiliary variable that serves as an upper bound on the objective. This way, a nonlinear objective can be minimized as long as the bound condition can be written as a conic constraint. Then, the linear objective to minimize is just the auxiliary variable itself. For example, the Shur complement can be used for the optimization of a quadratic objective.

The Shur complement allows to express a quadratic constraint as a semidefinite constraint.

2.3.1 Proposition (Shur Complement)**.** *Let M be a hermitian matrix partitioned as*

$$M = \begin{pmatrix} A & B \\ B^* & C \end{pmatrix},$$

where A and C are square. Then,

$$M > 0 \quad \Leftrightarrow \quad A > 0,\, C - B^* A^{-1} B > 0.$$

The matrix $C - B^* A^{-1} B$ is called Shur complement. The proof is interesting, because it only relies on basic properties of the semidefinite cone S:

2.3.2 Lemma. *Let M, X be $n \times n$ matrices, then*

$$M \geq 0 \quad \Leftrightarrow \quad \det X \neq 0,\, X^* M X \geq 0.$$

Proof. $M \geq 0$ is equivalent to the existence of a matrix B such that $M = B^* B$. Therefore $X^* M X = X^* B^* B C = (BX)^* (BX) \geq 0$. On the other hand, since X is invertible, for every vector z there exists a vector $y = X^{-1} z$ such that $z = Xy$. Hence $\langle z | Mz \rangle = \langle Xy | MXy \rangle = \langle y | X^* M X y \rangle \geq 0$. ∎

With this Lemma, the Proposition 2.3.1 about the Shur complement can easily be shown [45].

Proof (Prop. 2.3.1). Let

$$X = \begin{pmatrix} \mathbb{1} & -A^{-1} B \\ 0 & \mathbb{1} \end{pmatrix}.$$

Then, $detX = 1$ and $M > 0$ if and only if $X^* M X > 0$ with the above Lemma for strict inequality. As

$$X^* \begin{pmatrix} A & B \\ B^* & C \end{pmatrix} X = \begin{pmatrix} A & 0 \\ 0 & C - B^* A^{-1} B \end{pmatrix}$$

and a direct sum of matrices is positive if and only if the summands are positive, the result follows. ■

Most texts and software packages about convex optimization only consider the case, where the underlying vector space is $\mathbb{R}^n$. In principle, this includes the complex case, as the complex case can be reduced to the real case [46]. To express the cone of complex hermitian positive semidefinite matrices using real symmetric positive semidefinite matrices, consider the linear transformation $T : \mathbb{C}^{d\times d} \to \mathbb{R}^{2d\times 2d}$,

$$Tx = \begin{pmatrix} \operatorname{Re} x & -\operatorname{Im} x \\ \operatorname{Im} x & \operatorname{Re} x \end{pmatrix}, \tag{2.2}$$

where $\operatorname{Re} x$ denotes the real part and $\operatorname{Im} x$ denotes the imaginary part of x. If x is hermitian, $\operatorname{Re} x$ is symmetric, $\operatorname{Im} x$ is antisymmetric ($(\operatorname{Im} x)^\top = -\operatorname{Im} x^* = -\operatorname{Im} x$) and hence Tx is symmetric. On the other hand, a symmetric matrix with block structure

$$s = \begin{pmatrix} a & -b \\ b & a \end{pmatrix} \tag{2.3}$$

leads to a hermitian matrix $x = a + ib$.

The fact that a complex conic program formulated with real variables is again a conic program is based on the observation that T does not change the scalar product (up to a factor) and that the above block structure is a convex constraint. For any two hermitian positive semidefinite matrices x, y we have

$$\begin{aligned} \langle Tx|Ty\rangle &= \operatorname{tr}((Tx)^\top Ty) \\ &= \operatorname{tr}\left(\begin{pmatrix} \operatorname{Re} x & \operatorname{Im} x \\ -\operatorname{Im} x & \operatorname{Re} x \end{pmatrix}\begin{pmatrix} \operatorname{Re} y & -\operatorname{Im} y \\ \operatorname{Im} y & \Re y \end{pmatrix}\right) \\ &= 2\operatorname{tr}(\operatorname{Re} x \operatorname{Re} y + \operatorname{Im} x \operatorname{Im} y) = 2\langle x|y\rangle \end{aligned}$$

since

$$\begin{aligned} &\langle \operatorname{Re} x + i \operatorname{Im} x | \operatorname{Re} y + i \operatorname{Im} y\rangle \\ &= \operatorname{tr}\left(\operatorname{Re} x \operatorname{Re} y + \operatorname{Im} x \operatorname{Im} y + i \operatorname{Re} x \operatorname{Im} y - i \operatorname{Im} x \operatorname{Re} y\right) \end{aligned}$$

and the imaginary part vanishes as x and y are positive ($x = a^*a$, $y = b^*b$, $\operatorname{tr}(x^*y) = \operatorname{tr}((ba^*)^*ba^*) \geq 0$).

Now consider the block decomposition of a real symmetric matrix

$$s = \begin{pmatrix} a & b \\ c & d \end{pmatrix}.$$

To ensure the block structure (2.3), we need additional constraints. Let $|i\rangle\langle j|$ denote a matrix basis element in the computational basis. As s is symmetric, we already

have $a^\top = a$, $d^\top = d$. We only have to ensure that the upper or lower triangular parts of a and d are the same,

$$\langle \begin{pmatrix} |i\rangle\langle j| & 0 \\ 0 & -|i\rangle\langle j| \end{pmatrix} |s\rangle = 0, \quad i, j = 1, \ldots, d;\ i < j.$$

Symmetry also already implies that $b^\top = c$, so, again, we only have to ensure equality of upper or lower triangular part,

$$\langle \begin{pmatrix} 0 & |i\rangle\langle j| \\ |i\rangle\langle j| & 0 \end{pmatrix} |s\rangle = 0, \quad i, j = 1, \ldots, d;\ i < j.$$

These constraints are linear, and hence can be modeled in a conic program.

A software packages for the optimization of complex conic problems is given by Sturm [47], so the conversion into a real valued problem can be avoided.

Chapter 3

Quantum Error Correction

3.1 Introduction

Quantum error correction is one of the key technologies for quantum computation. Quantum effects are mostly absent in everyday life. So while scaling a quantum system, it is a nontrivial task to preserve the quantum nature that is responsible for the exponential speedup as in the Shor algorithm [48]. In a quantum computer, one has to fight the natural decoherence as well as unwanted interactions with the environment that introduce errors in the computation. Hence the task is to avoid, detect, and correct these quantum errors.

Quantum errors pose new challenges to correction algorithms, as the situation is fundamentally different from today's digital computers. Not only do we have an analog device, we also have new phenomenons such as entanglement, and the lack of a cloning possibility. Due to the No-Cloning Theorem [5], encoding based on simple redundancy and majority vote as decoding is not possible. Therefore, we have to look for more subtle ways to distribute quantum information among larger systems in order to be able to reduce errors. Quantum error correction schemes are based on increased distinguishability rather than on redundancy. Since the explicit constructions of such codes by Shor [49] and Steane [50], we know that such schemes exist. Furthermore, the theory of stabilizer codes[1] even provides us with a tool to create such encodings for larger systems. In this chapter, the focus is on fault tolerant storage and transmission of quantum systems rather than on fault tolerant computation.

The setting is as follows: Given an unknown quantum state ρ on a d-dimensional Hilbert space $\mathcal{H}$, the noise that acts on that state is described by a quantum channel $\widehat{T}\colon \mathcal{B}(\mathcal{H}) \to \mathcal{B}(\mathcal{H})$, $\rho \mapsto \widehat{T}(\rho)$. The idea is, given a larger quantum system $\mathcal{K}$ with noise $T\colon \mathcal{B}(\mathcal{K}) \to \mathcal{B}(\mathcal{K})$, to find a quantum code, such that the conjunction $E \circ T \circ D$

[1] See [51] or [35] for a survey on the existing techniques.

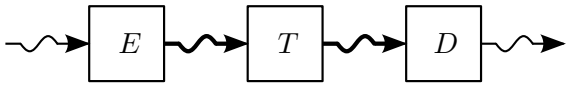

Figure 3.1: Quantum error correction setting. The encoder E encodes the state of a smaller quantum system into a state of the larger quantum system that is subject to the noise T. The decoder D then tries to recover the state of the smaller quantum system.

is closer to the ideal channel $\mathrm{id}\colon \rho \mapsto \rho$ than $\widehat{T}$ is.

3.1.1 Definition (Quantum Code)**.** A quantum code for the Hilbert spaces $\mathcal{H}$ and $\mathcal{K}$ is given by an encoding channel $E\colon \mathcal{B}(\mathcal{H}) \to \mathcal{B}(\mathcal{K})$ and a decoding channel $D\colon \mathcal{B}(\mathcal{K}) \to B(\mathcal{H})$.

The quantum error correction setting is shown in Figure 3.1. A common choice for $\mathcal{K}$ is the n-th tensor product of $\mathcal{H}$, $\mathcal{K} = \mathcal{H}^{\otimes n}$, with the assumption that $T = \widehat{T}^{\otimes n}$. However, this doesn't have to be the case and we will allow noisy interactions between tensor factors.

3.1.1 Perfect Correction

In the ideal case, perfect error correction is possible, that is, one can find encoding and decoding channels, such that

$$(E \circ T \circ D)(\rho) = \mathrm{id}(\rho) = \rho, \quad \forall \rho \in \mathcal{B}(\mathcal{H}). \tag{3.1}$$

Due to convexity of the state space, it is sufficient to look at the correction of Kraus operators t_α of the noisy channel for pure states,

$$T(|\psi\rangle\langle\psi|) = \sum_\alpha t_\alpha |\psi\rangle\langle\psi| t_\alpha^*.$$

Indeed, if the code corrects the operators t_α, it also corrects all linear combinations of them. Hence, if a code corrects the noisy channel T with Kraus operators t_α perfectly, it also perfectly corrects all other channels that have Kraus operators in the space spanned by the t_α. For this reason, the focus is on the correction of all operators from an operator space $\mathcal{E} \subset \mathcal{B}(\mathcal{H})$. The elements $e \in \mathcal{E}$ of such a space are called error operators, or just errors. A basis of $\mathcal{E}$ is called error basis. So, in the case of a finite dimensional Hilbert space we have the remarkable situation that the correction of a discrete set of basis errors guarantees the correction of a continuous set of errors.

A particular class of error bases provides the connection between quantum error correction and group theory [52].

3.1.2 Definition (Nice Error Basis)**.** A basis $\{e_g \,|g \in G\}$ of a linear operator space $\mathcal{E} \subset \mathcal{B}(\mathcal{H})$, $\dim \mathcal{H} = d$, is called *nice error basis*, if

1. e_g is a unitary operator on $\mathcal{H}$,
2. G is a group of order d^2,
3. $\mathrm{tr}(e_g) = d\delta_{g,1}$,
4. $e_g e_h = \omega_{g,h} e_{gh}$, $\omega_{g,h} \in \mathbb{C}$, $|\omega_{g,h}| = 1$, for all $g, h \in G$.

In the qubit case, a nice error basis is given by the Pauli operators:

identity	$I = \sigma_0 = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$,	$I\|a\rangle = \|a\rangle$
bit-flip	$X = \sigma_1 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$,	$X\|a\rangle = \|a \oplus 1\rangle$
phase-flip	$Z = \sigma_3 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$,	$Z\|a\rangle = (-1)^a\|a\rangle$
combined bit- and phase-flip	$Y = \sigma_2 = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}$,	$Y\|a\rangle = iXZ\|a\rangle$ $= i(-1)^a\|a \oplus 1\rangle$,

with $a \in \{0, 1\}$ and the mapping $|0\rangle \sim \begin{pmatrix} 1 \\ 0 \end{pmatrix}$, $|1\rangle \sim \begin{pmatrix} 0 \\ 1 \end{pmatrix}$.

For larger dimensions, nice error bases are given by discrete Weyl systems[2]. The connection to group theory lead to the development of so called stabilizer codes [9]. These codes are designed to perfectly correct errors that are tensor products of Pauli operators.

To do so, stabilizer codes use a special type of encoders, namely channels with a single Kraus operator.

3.1.3 Definition (Isometric Encoder)**.** An isometric encoder $E\colon \mathcal{B}(\mathcal{H}) \to \mathcal{B}(\mathcal{K})$ is a channel of the form $E(\rho) = v\rho v^*$, with an isometry v, i. e., $v^*v = \mathbb{1}$.

The columns of the isometry v can be interpreted as codewords for the computational basis of ρ. Thus, isometric encoding can be thought of as increasing the distinguishability of the basis states. For isometric encoders, there exist a general statement about their ability to correct errors, probably the most important insight into perfect error correction so far [10]:

[2] See [51] for more details.

3.1.4 Theorem (Knill-Laflamme)**.** *Given an operator space $\mathcal{E} \subset \mathcal{B}(\mathcal{K})$. Let $E(x) = v^*xv$ be an isometric encoder with the isometry $v\colon \mathcal{H} \to \mathcal{K}$. Then, there exists a decoder such that this code can perfectly correct all noisy channels with Kraus operators in $\mathcal{E}$ if and only if*

$$v^* e_\alpha^* e_\beta v = \lambda_{e_\alpha^* e_\beta} \mathbb{1}_{\mathcal{H}}, \quad \lambda_{e_\alpha^* e_\beta} \in \mathbb{C}, \tag{3.2}$$

for all $e_\alpha, e_\beta \in \mathcal{E}$.

In the proof of the theorem[3], a decoding channel is explicitly constructed. Such a decoder can always be chosen to be homomorphic [54].

3.1.5 Definition (Homomorphic Decoder)**.** A homomorphic decoder $D\colon \mathcal{B}(\mathcal{H}) \to \mathcal{B}(\mathcal{K})$ is a completely positive map with $D(\mathbb{1}) \leq \mathbb{1}$ and

$$D(xy) = D(x)D(y)$$

for all $x, y \in \mathcal{B}(\mathcal{H})$.

We generally refer to a completely positive map D with $D(\mathbb{1}) \leq \mathbb{1}$ as a subchannel. In contrast to channels, subchannels may produce no output result for some inputs. This is sometimes used to shorten the description of decoders for input cases that can never occur in a particular setting.

Note that if D is a channel, i. e., $D(\mathbb{1}) = \mathbb{1}$, then D has the form

$$D(x) = u(x \otimes \mathbb{1}_{\mathcal{D}})u^*,$$

with a dilation Hilbert space $\mathcal{D}$, where u is a unitary operator. The normalization condition implies $D(\mathbb{1}) = u(\mathbb{1} \otimes \mathbb{1}_{\mathcal{D}})u^* = uu^* = \mathbb{1}$, and $u^*u = \mathbb{1}$ follows from

$$D(xy) = u(xy \otimes \mathbb{1}_{\mathcal{D}})u^* = D(x)D(y) = u(x \otimes \mathbb{1}_{\mathcal{D}})u^*u(y \otimes \mathbb{1}_{\mathcal{D}})u^*.$$

With the Knill-Laflamme Theorem at hand and the ability to construct stabilizer codes, perfect error correction is possible for a certain type of errors: rare errors in multiple applications of the channel. Stabilizer codes are designed to correct error operators that are tensor factors of (usually only few) Pauli errors with the identity on all other factors. They are not designed to correct small errors, such as a small unitary overrotation. With limited resources, perfect error correction is often not possible, even for small errors. However, coding schemes can still be used to reduce the noise, which is the objective of approximate error correction.

3.1.2 Approximate Correction

In the case that no perfect error correction code exists for a given noise, it may still be possible to reduce the noise with an approximate error correcting code. So in the

[3] For example, see [53] for a proof of the Knill-Laflamme Theorem in the above formulation.

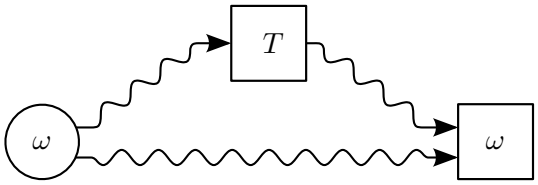

Figure 3.2: Definition of the channel fidelity for a channel T. The maximal entangled state ω is used as input of $T \otimes \mathrm{id}$, as well as observable.

situation where, for given $\mathcal{K}$, no solution (E, D) to (3.1) exists, we are looking for a code such that the resulting channel $E \circ T \circ D$ is more close to the ideal channel than the initial noise $\widehat{T}$. As a measure of how close a given channel T is to the ideal channel, we will use a special case of Schumacher's entanglement fidelity[4]:

3.1.6 Definition (Channel Fidelity)**.** Let $\mathcal{H}$ be a d-dimensional Hilbert space, $\omega = |\Omega\rangle\langle\Omega|$ the maximally entangled state with $|\Omega\rangle = \frac{1}{\sqrt{d}} \sum |ii\rangle \in \mathcal{H} \otimes \mathcal{H}$, and let id be the ideal channel on $\mathcal{B}(\mathcal{H})$. The *channel fidelity* of a channel $T\colon \mathcal{B}(\mathcal{H}) \to \mathcal{B}(\mathcal{H})$ is defined as

$$\begin{aligned} f_c(T) &= \mathrm{tr}(\omega(T \otimes \mathrm{id})(\omega)) \\ &= \langle\Omega|(T \otimes \mathrm{id})(|\Omega\rangle\langle\Omega|)|\Omega\rangle. \end{aligned} \tag{3.3}$$

The main virtue of choosing this fidelity as a figure characterizing the deviation from the identity is that it is linear in T. The definition is depicted in Figure 3.2. Note that due to the cyclicity of the trace, the channel fidelity has the same form in the Schrödinger picture,

$$f_c(T) = \mathrm{tr}(\omega(T \otimes \mathrm{id})(\omega)) = \mathrm{tr}((T_* \otimes \mathrm{id})(\omega)\omega) = \mathrm{tr}(\omega(T_* \otimes \mathrm{id})(\omega)) = f_c(T_*). \tag{3.4}$$

The channel fidelity has the following useful properties [53].

3.1.7 Proposition. *Let $\mathcal{H}$ be a d-dimensional Hilbert space, $T : \mathcal{B}(\mathcal{H}) \to \mathcal{B}(\mathcal{H})$ a channel.*

- *f_c is linear in T.*
- *$f_c(T)$ is continuous with respect to the operator norm.*
- *$0 \leq f_c(T) \leq 1$ for all channels T.*
- *$f_c(T) = 1 \quad \Leftrightarrow \quad T = id$.*

[4]The channel fidelity equals Schumacher's entanglement fidelity [55] for the choice of state $\rho = \mathrm{tr}_{\mathcal{H}}\, \omega = 1/d\mathbb{1}$, where ω is the maximal entangled state.

- *Given the Kraus operators t_α of T, $T(A) = \sum_\alpha t_\alpha^* A\, t_\alpha$, the channel fidelity can be written as*
$$f_c(T) = \frac{1}{d^2} \sum_\alpha |\operatorname{tr}(t_\alpha)|^2. \tag{3.5}$$

Furthermore, the channel fidelity is directly related to the mean fidelity for pure input states [56]. However, observe that for a channel $T(A) = \sigma_x A \sigma_x$, the channel fidelity is $f_c(T) = 0$, but perfect correction is possible simply by an additional rotation σ_x. The linearity of the channel fidelity makes it especially valuable as objective for convex optimization, as we will do later in this chapter. Such a linear criterion is possible only because the ideal channel is on the boundary of the set of channels. Linearity is the reason why we prefer the channel fidelity to the norm of completely boundedness (cb-norm),

$$\|T\|_{\mathrm{cb}} = \sup\Big\{ \|(T \otimes \mathrm{id}_n)(A)\| \Big| n \in \mathbb{N}; A \in \mathcal{B}(\mathcal{H} \otimes \mathbb{C}^n); \|A\| \leq 1 \Big\}.$$

As an example, we look at the fivefold tensor product of the qubit depolarizing channel as noise, $T_p \colon \mathcal{B}(\mathcal{H}) \to \mathcal{B}(\mathcal{H})$,

$$T_p(A) = p \operatorname{tr}\left(A \frac{1}{d} \mathbb{1} \right) \mathbb{1} + (1-p)\, A. \tag{3.6}$$

That is, we have $\dim \mathcal{H} = d = 2$, $\mathcal{K} = \mathcal{H}^{\otimes 5}$, and $T(A) = T_p^{\otimes 5}(A)$. If we decompose (3.6) into Pauli operators, we get

$$T_p(A) = \frac{p}{4}\Big(X A X + Y A Y + Z A Z \Big) + \Big(1 - \frac{3}{4} p\Big) \mathbb{1} A \mathbb{1}. \tag{3.7}$$

We can now apply the five bit code (E_5, D_5) [57, 58] as an example of a stabilizer code. The five bit code is designed to correct all Kraus operators of T_p^5 that have at most one tensor factor different from the matrix unity. Clearly, only few Kraus operators of T_p^5 are of this form, so perfect correction is not possible. However, from (3.7) we see that all terms of order p are corrected. A longer calculation [53] shows that the fidelity using the five bit code is

$$f_c(E_5 T_p^{\otimes 5} D_5) = 1 - \frac{45}{8} p^2 + \frac{75}{8} p^3 - \frac{45}{8} p^4 + \frac{9}{8} p^5 \tag{3.8}$$

compared to the fidelity

$$f_c(T_p) = 1 - \frac{3}{4} p. \tag{3.9}$$

of the single use of the depolarizing channel. That is, the five bit code reduces the error from order p to p^2. This connection between the number of errors that a code corrects perfectly and the approximate correction performance for an arbitrary channel is true more generally. See [51] for a quantitative statement in terms of the

cb-norm. Furthermore, isometric encoding is optimal with respect to the quantum capacity[5].

However, although codes for perfect error correction can be applied in the approximate correction setting, they can be outperformed by codes adapted to this setting. For example, in [14], Leung, Nielsen, Chuang, and Yamamoto construct such a code for a specific noisy channel that does not satisfy the Knill-Laflamme condition (3.2), but violates a bound for perfect correcting codes. Also, in [60], Crépeau, Gottesman, and Smith construct a family of approximate error correcting codes that correct more localized errors than it is possible to correct with a perfect error correcting code. Moreover, the fidelity obtained with their codes is exponentially good in the number of qubits used in the intermediate Hilbert space.

Another special case in approximate quantum error correction is the reversal of a quantum channel, i. e., $\mathcal{H} = \mathcal{K}$ and $E = \mathrm{id}$. In [61], Barnum and Knill give a construction scheme for an approximate reversal channel and show that the resulting fidelity is close to that of the (unknown) optimal reversal operation. A more general result regarding approximate error correction is given by Schumacher and Westmoreland [62]. They show that approximate error correction is possible, if the loss of coherent information is small. That is, they establish a connection between approximate correction and an entropic quantity.

3.2 Optimal Codes for Approximate Correction

Here, we make a more direct approach and treat approximate quantum error correction as the optimization problem

$$\max_{(E,D)} f_c(ETD). \tag{3.10}$$

This is particularly interesting in the low dimensional case with fixed resource $\mathcal{K}$, that is, the maximal Hilbert space dimension that can be engineered in the laboratory is fixed. Furthermore, we will make no assumptions about E, D or T. In particular, we will not assume that E is an isometric encoder or that D is a homomorphic decoder, nor will we assume a tensor structure of the noise T.

The optimization (3.10) is a joint optimization over the encoder E and the decoder D. We will solve this optimization by alternately optimizing over the encoder and decoder, that is, either E or D is changed to improve the fidelity in each step.

3.2.1 Definition (Seesaw Iteration for Error Correcting Codes)**.** Let f be a fidelity for channels $S\colon \mathcal{B}(\mathcal{H}) \to \mathcal{B}(\mathcal{H})$, that quantifies how close a given channel is to the

[5]Quantum capacity is the highest possible number of qubit transmissions per use of the channel using a suitable code and in the limit of large messages to transmit. See [59] for an overview.

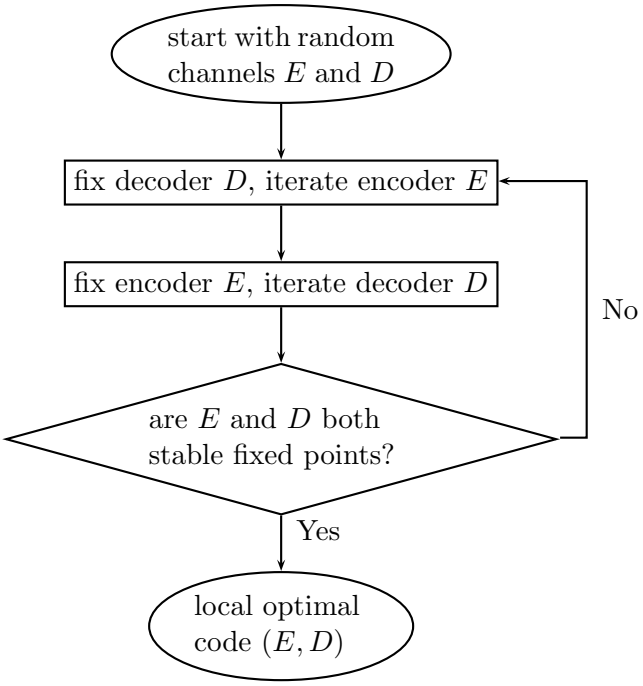

Figure 3.3: Seesaw iteration algorithm for an optimal error correcting code.

identity. Furthermore, let $T\colon \mathcal{B}(\mathcal{K}) \to \mathcal{B}(\mathcal{K})$ be the noisy channel. The *seesaw iteration* is then defined by:

1. Randomly choose a decoder channel $D_0\colon \mathcal{B}(\mathcal{H}) \to \mathcal{B}(\mathcal{K})$.
2. Choose the encoder channel $E_{n+1}\colon \mathcal{B}(\mathcal{K}) \to \mathcal{B}(\mathcal{H})$ as a channel that maximizes $E \mapsto f(ETD_n)$.
3. Choose the decoder channel $D_{n+1}\colon \mathcal{B}(\mathcal{H}) \to \mathcal{B}(\mathcal{K})$ as a channel that maximizes $D \mapsto f(E_{n+1}TD)$.
4. Exit, if E_{n+1} and D_{n+1} are both stable fixed points. Otherwise continue with step 2.

The algorithm is shown in Figure 3.3. It will result in a local optimum (E, D) and several runs from different starting points will be necessary to build up confidence that (E, D) is indeed a global optimum. Note that the seesaw iteration can also find error correcting codes for the implementation of reversible channels, most notably unitary gates. If ETD implements the unitary channel $U(A) = u^*Au$, then $uETD(\cdot)u^*$ is the ideal channel, and the objective to optimize becomes $f(uETD(\cdot)u^*)$.

Breaking up the joint optimization with the seesaw iteration reduces it to two single optimization problems of the following form:

3.2.2 Problem (Optimizing a Linear Functional over Channels)**.** Let $S\colon \mathcal{B}(\mathcal{H}_1) \to \mathcal{B}(\mathcal{H}_2)$ be a channel, where $\mathcal{H}_1$ and $\mathcal{H}_2$ are fixed Hilbert spaces of finite dimension. Given a linear objective f that is positive on all positive maps, the problem is to find a channel S that maximizes f,

$$f(S) = \max_T f(T). \tag{3.11}$$

Problem 3.2.2 is a semidefinite program [63] as we will verify below, that is, a linear optimization problem with a semidefinite constraint.

We will now have a closer look on the constraints of the optimization (3.11). In this optimization, $T\colon \mathcal{B}(\mathcal{H}_1) \to \mathcal{B}(\mathcal{H}_2)$ is constrained to be a channel. This means that the map T (in Heisenberg picture) has to be linear, completely positive, and unital. Equivalently, this means that the predual (Schrödinger picture) T_* has to be a linear, completely positive, and trace preserving map. By definition, the map T is completely positive if $T \otimes id_{\mathcal{K}}$ maps positive operators to positive operators for any Hilbert space $\mathcal{K}$. In order to be able to easily test for complete positivity, we make use of a one-to-one correspondence between channels and Hilbert space operators [64, 65].

3.2.3 Definition (Jamiolkowsky Dual of a Channel)**.** Let $\mathcal{H}_1$, $\mathcal{H}_2$ be finite dimensional Hilbert spaces and let $\mathcal{L}^2(\mathcal{H}_1, \mathcal{H}_2)$ denote the space of Hilbert Schmidt operators from $\mathcal{H}_1$ to $\mathcal{H}_2$ with scalar product $\langle\!\langle x|y\rangle\!\rangle = \operatorname{tr}(x^*y)$. Then, if $|i\rangle$, $i = 1, \ldots, \dim \mathcal{H}_1$ denotes the vectors of a basis of $\mathcal{H}_1$, we associate with any map $T\colon \mathcal{B}(\mathcal{H}_1) \to \mathcal{B}(\mathcal{H}_2)$ an operator $\widetilde{T} \in \mathcal{B}\big(\mathcal{L}^2(\mathcal{H}_1, \mathcal{H}_2)\big)$ by

$$\widetilde{T}(x) = \sum_{i,j} T(|i\rangle\langle j|)x|j\rangle\langle i| \tag{3.12}$$

with inversion formula

$$T(A) = \sum_{\mu,j} \widetilde{T}(|\mu\rangle\langle j|)A|j\rangle\langle\mu| \tag{3.13}$$

where $|\mu\rangle$ is a set of basis vectors for $\mathcal{H}_2$.

In short, the defining equation (3.12) can be written in the form

$$\langle a|T(|i\rangle\langle j|)|b\rangle = \langle a|\widetilde{T}(|b\rangle\langle j|)|i\rangle. \tag{3.14}$$

With the linear isomorphism $\mathcal{L}^2(\mathcal{H}_1, \mathcal{H}_2) \simeq \mathcal{H}_1 \otimes \overline{\mathcal{H}}_2$, $|\varphi\rangle\langle\psi| \simeq \varphi \otimes \overline{\psi}$, equation (3.12) can be seen as the definition of matrix elements of a matrix $\widehat{T}$,

$$\langle a\otimes\bar{i}|\widehat{T}(b\otimes\bar{j})\rangle = \langle\!\langle |a\rangle\langle i||\widetilde{T}(|b\rangle\langle j|)\rangle\!\rangle = \operatorname{tr}((|a\rangle\langle i|)^*\widetilde{T}(|b\rangle\langle j|)) = \langle a|\widetilde{T}(|b\rangle\langle j|)|i\rangle. \tag{3.15}$$

So in this sense, equation (3.14) amounts to a reshuffling of matrix elements.

The key feature of the correspondence between T and $\widetilde{T}$ is as follows:

3.2.4 Proposition. *Let T be a map $T\colon \mathcal{B}(\mathcal{H}_1) \to \mathcal{B}(\mathcal{H}_2)$. Then T is completely positive if and only if $\widetilde{T}$ is a positive definite operator on the Hilbert space $\mathcal{L}^2(\mathcal{H}_1, \mathcal{H}_2)$.*

Proof. Let T be completely positive. Then T has a Kraus decomposition $T(x) = \sum_\alpha t_\alpha^* x t_\alpha$ with operators $t_\alpha \in \mathcal{L}^2(\mathcal{H}_2, \mathcal{H}_1)$. We can therefore write equation (3.12) as

$$\widetilde{T}(x) = \sum_{ij} \sum_\alpha t_\alpha^* |i\rangle\langle j| t_\alpha x |j\rangle\langle i| = \sum_\alpha t_\alpha^* \operatorname{tr}(t_\alpha x).$$

Thus, we have

$$\big\langle\!\big\langle x \big| \widetilde{T}(x) \big\rangle\!\big\rangle = \sum_\alpha |\operatorname{tr}(t_\alpha x)|^2,$$

and we can write $\widetilde{T}$ as

$$\widetilde{T} = \sum_\alpha \big|t_\alpha^*\big\rangle\!\big\rangle\big\langle\!\big\langle t_\alpha^*\big|, \tag{3.16}$$

which implies that $\widetilde{T}$ is positive.

Conversely, let $\widetilde{T} \geq 0$. Then $\widetilde{T}$ has the decomposition $\widetilde{T} = \sum_\alpha \big|t_\alpha^*\big\rangle\!\big\rangle\big\langle\!\big\langle t_\alpha^*\big|$. With the inversion formula (3.13), we get

$$\begin{aligned} T(A) &= \sum_{\mu j} \sum_\alpha t_\alpha^* \operatorname{tr}(t_\alpha |\mu\rangle\langle j|) A |j\rangle\langle\mu| \\ &= \sum_{\mu j} \sum_\alpha t_\alpha^* A \langle j|t_\alpha|\mu\rangle\, |j\rangle\langle\mu| = \sum_\alpha t_\alpha^* A t_\alpha. \end{aligned}$$

Thus, T is a completely positive map. ∎

With the matrix representation (3.15) of $\widetilde{T}$, Proposition 3.2.4 reads:

$$\widehat{T} \geq 0 \Leftrightarrow T \text{ is completely positive.} \tag{3.17}$$

We will now express the remaining constraint, $T(\mathbb{1}) = \mathbb{1}$, in terms of $\widetilde{T}$, or equivalently, $\widehat{T}$:

3.2.5 Lemma.

1. *The map $T\colon \mathcal{B}(\mathcal{H}_1) \to \mathcal{B}(\mathcal{H}_2)$ is unital, i. e., $T(\mathbb{1}) = \mathbb{1}$, if and only if*

$$\operatorname{tr}_{\mathcal{H}_1} \widehat{T} = \mathbb{1}_{\mathcal{H}_2}, \tag{3.18}$$

   where $\operatorname{tr}_{\mathcal{H}_1}$ denotes the partial trace over the Hilbert space $\mathcal{H}_1$.

2. *The map $T_*\colon \mathcal{B}(\mathcal{H}_1) \to \mathcal{B}(\mathcal{H}_2)$ is trace preserving, i. e., $\operatorname{tr} T_*(\rho) = \operatorname{tr} \rho$, if and only if*

$$\operatorname{tr}_{\mathcal{H}_2} \widehat{T}_* = \mathbb{1}_{\mathcal{H}_1}, \tag{3.19}$$

   where $\operatorname{tr}_{\mathcal{H}_2}$ denotes the partial trace over the Hilbert space $\mathcal{H}_2$.

Proof. The first part follows from the equality

$$\begin{aligned}\langle a|\mathbb{1}|b\rangle &= \sum_i \langle a|T(|i\rangle\langle i|)|b\rangle = \sum_i \langle a|\widetilde{T}(|b\rangle\langle i|)|i\rangle \\ &= \sum_i \langle a\otimes\bar{i}|\widehat{T}(b\otimes\bar{i})\rangle = \langle a|\operatorname{tr}_{\mathcal{H}_1}(\widehat{T})|b\rangle.\end{aligned}$$

For the second part, we write the trace preservation condition as $\operatorname{tr} T_*(|i\rangle\langle j|) = \delta_{ij}$ for all basis vectors i, j and conclude that

$$\langle i|\mathbb{1}|j\rangle = \sum_a \langle a\otimes\bar{i}|\widehat{T}_*(a\otimes\bar{j})\rangle = \langle i|\operatorname{tr}_{\mathcal{H}_2}(\widehat{T}_*)|j\rangle.$$

■

As we have a positive objective, we can also allow subchannels, i. e., completely positive linear maps with $T(\mathbb{1}) \leq \mathbb{1}$, in the maximization (3.11). If the optimum is attained on a subchannel T, we can always add Kraus operators, such that the equality holds, without lowering the objective. In terms of $\widehat{T}$, the subchannel constraint can be written in Heisenberg picture as

$$\mathbb{1} - \operatorname{tr}_{\mathcal{H}_1} \widehat{T} \geq 0. \tag{3.20}$$

Combining the channel constraints (3.17) and (3.20) we get the semidefinite constraint

$$\widehat{T} \oplus (\mathbb{1} - \operatorname{tr}_{\mathcal{H}_1} \widehat{T}) \geq 0, \tag{3.21}$$

so Problem 3.2.2 is indeed a semidefinite program.

We will now show that every optimization over subchannels of a linear positive functional can be written in terms of the channel fidelity and the Jamiolkowsky dual. So consider the positive linear functional f to optimize over channels $T\colon \mathcal{B}(\mathcal{H}_1) \to \mathcal{B}(\mathcal{H}_2)$. Applying Riesz Theorem [42], we can write every continuous linear functional of $\widetilde{T}$ as scalar product with another operator, say $\widetilde{F}_*$, $f(\widetilde{T}) = \langle\!\langle \widetilde{F}_* | \widetilde{T} \rangle\!\rangle = \operatorname{tr}(\widetilde{F}\widetilde{T})$. As f is positive, $\widetilde{F}$ is a positive operator. With the Kraus operators t_α of T and equation (3.16) we get

$$\begin{aligned}f(T) = \operatorname{tr}(\widetilde{F}\widetilde{T}) &= \sum_\alpha \operatorname{tr}(\widetilde{F}|t_\alpha^*\rangle\!\rangle\langle\!\langle t_\alpha^*|) \\ &= \sum_\alpha \langle\!\langle t_\alpha^*|\widetilde{F}|t_\alpha^*\rangle\!\rangle \\ &= \sum_\alpha \operatorname{tr}(t_\alpha \widetilde{F}(t_\alpha^*)),\end{aligned} \tag{3.22}$$

where we used the definition of the Hilbert Schmidt scalar product $\langle\!\langle x|y\rangle\!\rangle = \operatorname{tr}(x^* y)$

in the last equation. Inserting the definition (3.12) for $\widetilde{F}$, we get

$$\begin{aligned}\sum_\alpha \operatorname{tr}(t_\alpha \widetilde{F}(t_\alpha^*)) &= \sum_\alpha \operatorname{tr}\left(t_\alpha \sum_{ij} F(|i\rangle\langle j|) t_\alpha^* |j\rangle\langle i|\right) \\ &= \sum_{ij} \operatorname{tr}(F(|i\rangle\langle j|) T(|j\rangle\langle i|)).\end{aligned} \tag{3.23}$$

Note that F is completely positive, but not necessarily a channel, as the unital condition does not hold in general. Using the predual F_* of F, defined by $\operatorname{tr}(x^* F_*(y)) = \operatorname{tr}(F(x)^* y)$, and using the fact that for completely positive maps $F(x)^* = \sum_\alpha (f_\alpha^* x f_\alpha)^* = F(x^*)$, we obtain

$$\begin{aligned}\sum_{ij} \operatorname{tr}(F(|i\rangle\langle j|) T(|j\rangle\langle i|)) &= \sum_{ij} \operatorname{tr}(F(|j\rangle\langle i|)^* T(|j\rangle\langle i|)) \\ &= \sum_{ij} \operatorname{tr}(|i\rangle\langle j| F_*(T(|j\rangle\langle i|))) \\ &= \sum_{ij} \langle j| F_* T(|j\rangle\langle i|) |i\rangle.\end{aligned} \tag{3.24}$$

Rewriting the last expression in terms of the maximally entangled state $|\Omega\rangle = (\dim \mathcal{H}_1)^{-1/2} \sum_k |kk\rangle$ leads us to

$$\begin{aligned}\sum_{ij} \langle j| F_* T(|j\rangle\langle i|) |i\rangle &= \sum_{ijkl} \langle j|l\rangle\langle i|k\rangle \langle j| F_* T(|l\rangle\langle i|) |k\rangle \\ &= \sum_{ijkl} \langle ii| (\mathrm{id} \otimes F_* T(|ll\rangle\langle ii|) |kk\rangle \\ &= (\dim \mathcal{H}_1)^2 \langle \Omega| (\mathrm{id} \otimes F_* T)(|\Omega\rangle\langle\Omega|) |\Omega\rangle \\ &= (\dim \mathcal{H}_1)^2 f_c(F_* T),\end{aligned} \tag{3.25}$$

where we used equation (3.3) for the channel fidelity f_c. So we can write any linear positive objective of a channel T in terms of the channel fidelity f_c and a completely positive map F.

3.2.6 Proposition. *Let $f(S)$ be a linear positive functional for channels $S\colon \mathcal{B}(\mathcal{H}_1) \to \mathcal{B}(\mathcal{H}_2)$ with finite dimensional Hilbert spaces $\mathcal{H}_1$, $\mathcal{H}_2$. Then there exists a completely positive map $F\colon \mathcal{B}(\mathcal{H}_1) \to \mathcal{B}(\mathcal{H}_2)$, such that*

$$f(S) = (\dim \mathcal{H}_1)^2 f_c(F_* S), \tag{3.26}$$

where f_c denotes the channel fidelity from Definition 3.1.6.

Note that F is in general not normalized, i. e., $F(\mathbb{1}) \neq \mathbb{1}$. Furthermore, as $\dim \mathcal{H}_1$ is fixed, we can safely omit the factor in the optimization if convenient.

We will now determine the maps F and S corresponding to the iteration steps of the seesaw iteration algorithm in Figure 3.3. For the decoder optimization, we have the linear functional

$$D \mapsto f_c(ETD) = f_c(F_* S). \tag{3.27}$$

From this, we can immediately deduce that

$$F = (ET)_*, \quad S = D. \tag{3.28}$$

For the encoder optimization, $E \mapsto f_c(ETD)$, we rewrite the objective in the Schrödinger picture according to (3.4),

$$E_* \mapsto f_c(D_* T_* E_*) = f_c(F_* S). \tag{3.29}$$

This leads to the identification

$$F = TD, \quad S = E_*. \tag{3.30}$$

Observe that the channel fidelity $f_c(ETD)$ can also be regarded as a linear objective for T. If we can find a channel T_1, such that $f_c(ET_1D) = 1$, we know from Proposition 3.1.7 that ET_1D is the ideal channel, and hence, that (E, D) is a perfect error correcting code. In particular, we don't need a seesaw like iteration as for the optimal code (E, D), so this optimization can be done by a single convex optimization. Thus, we have a numerical algorithm that can build up confidence whether a particular code is a perfect error correcting code, possibly for a different type of noisy channel than it is designed to correct. We will now identify $\widetilde{F}$ for the corresponding objective

$$T \mapsto f_c(ETD). \tag{3.31}$$

Analog to (3.25), we can rewrite the channel fidelity in the form

$$\begin{aligned}(\dim \mathcal{H}_1)^2 f_c(ETD) &= \sum_{ab} \langle a|E(T(D(|a\rangle\langle b|)))|b\rangle \\ &= \sum_{ab}\sum_{\mu j} \langle a|E\Big(\widetilde{T}(|\mu\rangle\langle j|)\, D(|a\rangle\langle b|)|j\rangle\langle \mu|\Big)|b\rangle.\end{aligned} \tag{3.32}$$

Furthermore, from linearity we get

$$\begin{aligned}f(\widetilde{T}) &= \sum_{\alpha\beta} f(\big|\alpha\big\rangle\!\big\rangle\big\langle\!\big\langle\alpha\big|\widetilde{T}\big|\beta\big\rangle\!\big\rangle\big\langle\!\big\langle\beta\big|) = \sum_{\alpha\beta} \big\langle\!\big\langle\alpha\big|\widetilde{T}\big|\beta\big\rangle\!\big\rangle f(\big|\alpha\big\rangle\!\big\rangle\big\langle\!\big\langle\beta\big|) \\ &= \sum_{\alpha\beta} \big\langle\!\big\langle\alpha\big|\widetilde{T}\big|\beta\big\rangle\!\big\rangle\big\langle\!\big\langle\beta\big|\widetilde{F}\big|\alpha\big\rangle\!\big\rangle = \mathrm{tr}(\widetilde{T}\widetilde{F}).\end{aligned} \tag{3.33}$$

And thus, with equation (3.32) and in abuse of notation, we conclude that

$$\begin{aligned}\big\langle\!\big\langle\beta\big|\widetilde{F}\big|\alpha\big\rangle\!\big\rangle &= f(\big|\alpha\big\rangle\!\big\rangle\big\langle\!\big\langle\beta\big|) \\ &= \sum_{ab}\sum_{\mu j} \langle a|E\Big(\big|\alpha\big\rangle\!\big\rangle\big\langle\!\big\langle\beta\big|(|\mu\rangle\langle j|)\, D(|a\rangle\langle b|)|j\rangle\langle\mu|\Big)|b\rangle \\ &= \sum_{ab} \langle a|E\Big(\big|\alpha\big\rangle\!\big\rangle D(|a\rangle\langle b|)\big|\beta^*\big\rangle\!\big\rangle\Big)|b\rangle.\end{aligned} \tag{3.34}$$

In summary, we have seen that the encoder and decoder optimization in the seesaw algorithm of Figure 3.3 can be stated in terms of F_* and S, with moderate computational overhead given by the concatenation and adjoining in (3.28) and (3.30). Furthermore, we have the ability to test the perfect correction abilities of a code with the noisy channel optimization (3.31).

3.2.1 Channel Power Iteration

In the following, we will develop an iteration algorithm [1] to obtain a numerical solution for the single iteration steps (3.11), $f(S) = \max_T f(T)$. The key equation for this iteration is given by (3.22),

$$f(S) = \sum_\alpha \operatorname{tr}\left(s_\alpha \widetilde{F}(s_\alpha^*)\right). \tag{3.35}$$

In the special case of optimizing isometric encoders, $S(x) = v^* x v$, I showed in my diploma thesis [53] that this equation leads to the fixed point equation

$$\widetilde{F}(v) = v|\widetilde{F}(v)| \tag{3.36}$$

for an optimal isometry v. Consequently, the numerical optimization was done by calculating the encoder isometry of step $(n+1)$, v_{n+1}, as polar isometry of $\widetilde{F}(v_n)$. However, that this iteration leads to the global maximum was only shown for special cases. Furthermore, the seesaw iteration in Figure 3.3 requires the optimization of channels with multiple Kraus operators.

We will now extend the iteration algorithm for isometric encoding such that it can be used for the optimization of general channels. In particular, it will therefore allow to optimize a decoder channel. Note that the above iteration implicitly complies with the channel constraints, as for every isometry $v^* v = \mathbb{1}$. For general channels, the iteration must conform to the subchannel constraints (3.18),

$$\operatorname{tr}_{\mathcal{H}_1} \widetilde{T} \leq \mathbb{1}, \tag{3.37}$$

and (3.19),

$$\operatorname{tr}_{\mathcal{H}_2} \widetilde{T}_* \leq \mathbb{1}, \tag{3.38}$$

for Schrödinger and Heisenberg picture, respectively, where we identify $\widetilde{T}$ with its matrix representation $\widehat{T}$. Note that these constraints are operator inequalities, and do not merely fix a normalization factor. It is instructive, however, to compare the problem with a simplified one, in which only the trace of the subchannel inequality is imposed, i. e., in which we only demand that

$$\operatorname{tr} S(\mathbb{1}) = \sum_\alpha \left\langle\!\left\langle s_\alpha^* \middle| s_\alpha^* \right\rangle\!\right\rangle \leq (\dim \mathcal{H}_2)^2. \tag{3.39}$$

Up to a factor this is just saying that $\widetilde{S}$ is a density operator. With this relaxed constraint, the optimal value of equation (3.35) is the largest eigenvalue of $\widetilde{F}$, and it is attained for an S with a single Kraus summand given by a corresponding eigenvector.

A numerical algorithm to compute the eigenvector of the largest eigenvalue of a positive operator $\widetilde{F}$ is given by the power iteration [66, 67]. The power iteration

starts with a generic initial vector and iteratively applies the operator $\widetilde{F}$ on it. Then, all components belonging to other eigenspaces will be suppressed and we will end up asymptotically in the eigenspace of the largest eigenvalue. In order to avoid numerical problems with arithmetic precision, it is convenient to normalize the iterates in every step, so that we get an iteration on unit vectors. Otherwise, roundoff errors may be scaled by large components of the iterated vector. So in our case, every iteration step of the power iteration, that optimizes (3.35) with the simplified constraint, consists of

1. Apply $\widetilde{F}$ to the adjoint of every Kraus operator of S.
2. Normalize.

So assume $\widetilde{F}$ has the eigenvalue decomposition $\widetilde{F} = \sum_{i=1}^{n} \lambda_i |i\rangle\!\rangle\langle\!\langle i|$, with eigenvalues

$$0 \neq \lambda_1 > \lambda_2 \geq \ldots \geq \lambda_n \geq 0.$$

When we start with the vector $|s^{(0)}\rangle\!\rangle = \sum_{i=1}^{n} \mu_i^{(0)} |i\rangle\!\rangle$, $\mu_1^{(0)} \neq 0$, then the m-th step of the power iteration leads to

$$\begin{aligned} |s^{(m)}\rangle\!\rangle &= \frac{1}{\|\widetilde{F}^m |s^{(0)}\rangle\!\rangle\|} \widetilde{F}^m |s^{(0)}\rangle\!\rangle \\ &= \lambda_1^m \left(\mu_1^{(m)} |1\rangle\!\rangle + \sum_{i=2}^{n} \left(\frac{\lambda_i}{\lambda_1} \right)^m \mu_i^{(m)} |i\rangle\!\rangle \right) \xrightarrow{m\to\infty} |1\rangle\!\rangle. \end{aligned}$$

Here the coefficients $\mu_i^{(m)}$ are given by the normalization and the $\mu_i^{(0)}$ of the starting vector. Thus we have $\widetilde{F}|s^{(\infty)}\rangle\!\rangle = \lambda_1 |s^{(\infty)}\rangle\!\rangle$ with a linear rate of convergence, depending on λ_2/λ_1.

As we will show below, all we need to do to get an iteration that implies the subchannel constraint (3.37) or (3.38) is to change the meaning of "normalize" in this scheme. Note that there is nothing special about the Kraus decomposition as applying $\widetilde{F}$ to every Kraus operator of S is the same as

$$\widetilde{S}'_{n+1} = \widetilde{F} \widetilde{S}_n \widetilde{F}, \tag{3.40}$$

where $\widetilde{S}_n$ is the normalized $\widetilde{S}$ of the last iteration step and $\widetilde{S}'_{n+1}$ is the unnormalized $\widetilde{S}$ of the next step.

3.2.7 Definition. Let $S: \mathcal{B}(\mathcal{H}_1) \to \mathcal{B}(\mathcal{H}_2)$ be a completely positive map. Then we define the *normalized* version of S as

$$N(S)(x) = S(\mathbb{1})^{-\frac{1}{2}} S(x) S(\mathbb{1})^{-\frac{1}{2}}, \tag{3.41}$$

where $S(\mathbb{1})^{-\frac{1}{2}}$ denotes the pseudo-inverse of $S(\mathbb{1})^{\frac{1}{2}}$.

It is useful to express this in terms of the Kraus operators of S. So let $\{s_\alpha\}$ be the Kraus operators, $S(\mathbb{1}) = \sum_\alpha s_\alpha^* s_\alpha$. To compute the pseudo-inverse, we have to split the Hilbert space $\mathcal{H}_2$ into an orthogonal sum of the support of $S(\mathbb{1})$ and its kernel, i. e., the space of all vectors ϕ such that $s_\alpha \phi = 0$ for all α. On the support of $S(\mathbb{1})$ this operator is invertible, so we can simply compute the inverse in the functional calculus. Another way to describe the construction (3.41) is to set[6]

$$\begin{aligned} N(S)(x) &= \lim_{\varepsilon \to 0} h_\varepsilon^2 S(x) h_\varepsilon^2, \\ h_\varepsilon &= (S(\mathbb{1}) + \varepsilon \mathbb{1})^{-1/4}. \end{aligned} \tag{3.42}$$

Note that $N(S(\mathbb{1}))$ is the projection onto the support of $S(\mathbb{1})$, so $N(S)$ is a subchannel in general. It is a proper channel if and only if this projection is the identity, i. e., if $S(\mathbb{1})$ was invertible in the first place. Furthermore, in the case of encoder iteration (3.30), we have the identification $S = E_*$, i. e., S is a channel in Schrödinger picture. Thus, the normalization condition becomes $N_*(S)(x) = N(S^*)(x)$. In terms of Kraus operators, this means that

$$s_\alpha^* = e_\alpha \mapsto e_\alpha \left(\sum_\beta e_\beta^* e_\beta \right)^{-1/2} = s_\alpha^* \left(\sum_\beta s_\beta s_\beta^* \right)^{-1/2}.$$

With the normalization step of Definition 3.2.7, we have the following iteration.

3.2.8 Definition (Subchannel Power Iteration)**.** Let $\mathcal{H}_1$, $\mathcal{H}_2$ be finite dimensional Hilbert spaces. Furthermore, let f be a linear positive objective on channels $S \colon \mathcal{B}(\mathcal{H}_1) \to \mathcal{B}(\mathcal{H}_2)$, given by the positive operator $\widetilde{F} \in \mathcal{B}(\mathcal{L}^2(\mathcal{H}_1, \mathcal{H}_2))$ according to equation (3.35). Then, the *subchannel power iteration* is given by the iteration step $S_n \mapsto S_{n+1}$:

$$\widetilde{S}'_{n+1} = \widetilde{F} \widetilde{S}_n \widetilde{F}, \tag{3.43}$$

$$S_{n+1} = N(S'_{n+1}), \tag{3.44}$$

where $N(S)$ denotes the proper normalization depending on whether S is a channel in the Heisenberg picture or in the Schrödinger picture.

Note that the iteration does not change the number of Kraus operators. Therefore, the starting channel S_0 should contain the maximal number of independent Kraus operators, or equivalently, use the maximal minimal dilation dimension in Stinespring representation. On the other hand, if we restrict the encoder iteration (3.30) to a single Kraus operator v, we obtain

$$v_{n+1} = s_{1,n+1}^* = \widetilde{F}(e_{1,n}^*) \left(\widetilde{F}(e_{1,n}^*)^* \widetilde{F}(e_{1,n}^*) \right)^{-1/2} = \widetilde{F}(v_n) \left(\widetilde{F}(v_n)^* \widetilde{F}(v_n) \right)^{-1/2},$$

[6] The odd choice of powers will be convenient later.

which is the polar isometry of $\widetilde{F}(v)$, since

$$v_{n+1}|\widetilde{F}(v_n)| = \widetilde{F}(v_n)\left(\widetilde{F}(v_n)^*\widetilde{F}(v_n)\right)^{-1/2}\cdot\left(\widetilde{F}(v_n)^*\widetilde{F}(v_n)\right)^{1/2} = \widetilde{F}(v_n).$$

So the iteration for isometric encodings [53] is indeed a special case of the subchannel power iteration.

The main result for the subchannel power iteration is given by the following Theorem.

3.2.9 Theorem (Subchannel Power Iteration)**.** *The subchannel power iteration, given in Definition 3.2.8, is monotone. That is, for successive iterates S_n, S_{n+1} we have $f(S_{n+1}) \geq f(S_n)$. Furthermore, if $S_n(\mathbb{1})$ is invertible in all iteration steps, then every channel S for which the maximum of f is attained is a stable fixed point of the iteration.*

We will proof these properties below and drop the assumption about the existence of the inverse of $S_n(\mathbb{1})$ by extending the algorithm to ensure that this is the case near an optimal solution.

3.2.1.1 Monotonicity

It is instructive to once again look at the case of the search for the largest eigenvalue of a positive operator A, by iterating A on a vector ϕ and renormalizing it in every step, so that $\phi \mapsto \langle\phi|A^2\phi\rangle^{-1/2}A\phi$. The monotonicity of this iteration follows from the following Lemma, which will also be used in the proof of the monotonicity of the subchannel power iteration.

3.2.10 Lemma (Monotonicity)**.** *Let A be a bounded positive operator on a Hilbert space, and ϕ an arbitrary vector in that space. Then, for every vector $\phi \neq 0$,*

$$\langle\phi|A^3\phi\rangle \geq \frac{\langle\phi|A^2\phi\rangle\langle\phi|A\phi\rangle}{\langle\phi|\phi\rangle}\,. \tag{3.45}$$

Equality holds if and only if ϕ is an eigenvector of A.

More precisely, assume that in the inequality

$$\langle\phi|A^3\phi\rangle - \frac{\langle\phi|A^2\phi\rangle\langle\phi|A\phi\rangle}{\langle\phi|\phi\rangle} \leq \varepsilon\,. \tag{3.46}$$

Then, with $a = \|\phi\|^{-2}\langle\phi|A\phi\rangle$, we have

$$\|(A - a\mathbb{1})\phi\|^2 \leq \frac{\varepsilon}{a}\,. \tag{3.47}$$

Proof. Let A be bounded and positive, and denote by $M_k = \langle\phi|A^k\phi\rangle$,$k = 0, 1, \ldots$, the moments of the spectral measure in the vector ϕ. Then, for any $a, b \in \mathbb{R}, c \geq 0$ the operator

$$A(A - b\mathbb{1})^2 + c(A - a\mathbb{1})^2$$

is positive, which implies

$$\langle\phi|A(A-b\mathbb{1})^2|\phi\rangle + c\langle\phi|(A-a\mathbb{1})^2|\phi\rangle \geq 0 \tag{3.48}$$

$$\Leftrightarrow \quad (M_3 - 2bM_2 + b^2M_1) + c(M_2 - 2aM_1 + a^2M_0) \geq 0\,. \tag{3.49}$$

With $b = c = M_2/M_1$ and $a = M_1/M_0$, this becomes

$$M_3 - \frac{M_2^2}{M_1} + \frac{M_2}{M_1}\left(M_2 + \frac{M_1^2}{M_0}\right) \geq 0 \quad \Leftrightarrow \quad M_3 - \frac{M_1^2 M_2}{M_1 M_0} \geq 0,$$

which leads to the desired inequality (3.45),

$$M_3 \geq \frac{M_2 M_1}{M_0}. \tag{3.50}$$

Now suppose equation (3.46) holds for some $\varepsilon \geq 0$. Then the second term in (3.49) is also bounded by ε, as both terms in (3.48) are positive. Thus we have

$$\begin{aligned}\|(A-a\mathbb{1})\phi\|^2 &= (M_2 - 2aM_1 + a^2M_0)\\ &\leq \frac{\varepsilon}{c} = \varepsilon\frac{M_1}{M_2} \leq \varepsilon\frac{M_0}{M_1}\\ &= \frac{\varepsilon}{a},\end{aligned}$$

where we used that $M_2 \geq M_1^2/M_0$. For $\varepsilon = 0$ this means that ϕ is an eigenvector with eigenvalue a. ■

For the iteration of the largest eigenvalue, this Lemma immediately implies the monotonicity property

$$\frac{\langle A\phi|A|A\phi\rangle}{\langle A\phi|A\phi\rangle} \geq \frac{\langle\phi|A|\phi\rangle}{\langle\phi|\phi\rangle}. \tag{3.51}$$

Now consider the iteration on subchannels. We will apply the Lemma to an operator on the Hilbert space, which is the direct sum of copies of $\mathcal{L}^2(\mathcal{H}_1, \mathcal{H}_2)$, with one component for each Kraus operator of S. The vectors in this space are hence tuples $\psi = (\psi_1, \ldots, \psi_n)$, where N is the number of Kraus operators and $\psi_\alpha \in \mathcal{L}^2(\mathcal{H}_1, \mathcal{H}_2)$. The scalar product of this Hilbert space is given by

$$\langle\!\langle \psi | \phi \rangle\!\rangle = \sum_{\alpha=1}^{N} \mathrm{tr}(\psi_\alpha^* \phi_\alpha).$$

Then, we fix $\varepsilon > 0$ in (3.42) and take

$$\begin{aligned}(A\psi)_\alpha &= h_\varepsilon \widetilde{F}(h_\varepsilon \psi_\alpha)\\ \phi_\alpha &= h_\varepsilon^{-1} s_\alpha^*,\end{aligned}$$

with

$$\begin{aligned} h_\varepsilon &= (M+\varepsilon\mathbb{1})^{-1/4}, \\ M = S'_{n+1}(\mathbb{1}) &= \sum_\alpha \widetilde{F}(s^*_{\alpha,n})\widetilde{F}(s^*_{\alpha,n})^*. \end{aligned} \tag{3.52}$$

This implies that A is a positive operator, because $\langle\!\langle \psi | A\psi \rangle\!\rangle = \sum_\alpha \mathrm{tr}(\psi^*_\alpha h_\varepsilon \widetilde{F}(h_\varepsilon \psi_\alpha))$, which is positive by positivity of $\widetilde{F}$ and hermiticity of h_ε.

Note that, with the above choices, we have $(A\phi)_\alpha = h_\varepsilon \widetilde{F}(s^*_\alpha)$, and that

$$s^*_{\alpha,n+1} = \lim_{\varepsilon\to 0} h^2_\varepsilon \widetilde{F}(s^*_{\alpha,n}) \tag{3.53}$$

are the normalized Kraus operators of the next iteration step according to equation (3.42).

The successive moments of A appearing in the Lemma become

$$\begin{aligned} \langle\!\langle \phi | \phi \rangle\!\rangle &= \mathrm{tr}\left(h_\varepsilon^{-2}\sum_\alpha s^*_\alpha s_\alpha\right) \\ &= \mathrm{tr}\left((M+\varepsilon\mathbb{1})^{1/2}\sum_\alpha s^*_\alpha s_\alpha\right) \\ \langle\!\langle \phi | A\phi \rangle\!\rangle &= \sum_\alpha \mathrm{tr}\left(s_\alpha \widetilde{F}(s^*_\alpha)\right) \\ &= f(S_n) \\ \langle\!\langle \phi | A^2\phi \rangle\!\rangle &= \langle\!\langle A\phi | A\phi \rangle\!\rangle \\ &= \sum_\alpha \mathrm{tr}\left(\left(h_\varepsilon \widetilde{F}(s^*_\alpha)\right)^* h_\varepsilon \widetilde{F}(s^*_\alpha)\right) \\ &= \mathrm{tr}\left(h^2_\varepsilon \sum_\alpha \widetilde{F}(s^*_\alpha)\widetilde{F}(s^*_\alpha)^*\right) \\ &= \mathrm{tr}\left((M+\varepsilon\mathbb{1})^{-1/2} M\right) \\ \langle\!\langle \phi | A^3\phi \rangle\!\rangle &= \langle\!\langle A\phi | A | A\phi \rangle\!\rangle \\ &= \sum_\alpha \mathrm{tr}\left((h_\varepsilon \widetilde{F}(s^*_\alpha))^* h_\varepsilon \widetilde{F}(h^2_\varepsilon \widetilde{F}(s^*_\alpha))\right) \\ &= \sum_\alpha \mathrm{tr}\left((h^2_\varepsilon \widetilde{F}(s^*_\alpha))^* \widetilde{F}(h^2_\varepsilon \widetilde{F}(s^*_\alpha))\right) \\ &= f(S_{\varepsilon,n+1}), \end{aligned}$$

where $S_{\varepsilon,n+1}$ is the completely positive map with Kraus operators $s^*_{\alpha,\varepsilon,n+1} = h^2_\varepsilon \widetilde{F}(s^*_\alpha)$, which in the limit $\varepsilon \to 0$ goes to the next iterate. In this limit, inequality (3.45) of the monotonicity Lemma 3.2.10 becomes

$$f(S_{n+1}) \geq \frac{\mathrm{tr}\, M^{1/2}}{\mathrm{tr}(M^{1/2}\sum_\alpha s^*_\alpha s_\alpha)} f(S_n) \geq f(S_n), \tag{3.54}$$

where in the last step we used that $\sum_\alpha s_\alpha^* s_\alpha = S_n(\mathbb{1}) \leq \mathbb{1}$. This proves monotonicity of the subchannel power iteration for channels in Heisenberg picture.

For the iteration in the Schrödinger picture, i. e., iteration of $S_*(\cdot) = \sum_\alpha s_\alpha \cdot s_\alpha^*$ instead of S, as in the encoder iteration (3.30), the analogous analysis applies with the choice

$$\begin{aligned} (A\psi)_\alpha &= \widetilde{F}(\psi_\alpha g_\varepsilon) g_\varepsilon, \\ g_\varepsilon &= (M + \varepsilon \mathbb{1})^{-1/4}, \\ M &= \sum_\alpha \widetilde{F}(s_\alpha)^* \widetilde{F}(s_\alpha), \\ \phi_\alpha &= s_\alpha g_\varepsilon^{-1}. \end{aligned} \tag{3.55}$$

In this case, we have $\widetilde{S}_* = |s_\alpha\rangle\!\rangle\langle\!\langle s_\alpha|$, so from equation (3.22) we know that the objective is given by

$$f(S_*) = \sum_\alpha \operatorname{tr}\left(s_\alpha^* \widetilde{F}(s_\alpha)\right).$$

The Kraus operators of the next iteration step are given by

$$s_{\alpha,n+1} = \lim_{\varepsilon \to 0} \widetilde{F}(s_{\alpha,n}) g_\varepsilon^2.$$

With the above choice, the scalar products of Lemma 3.2.10 become

$$\begin{aligned} \langle\!\langle \Phi | \Phi \rangle\!\rangle &= \operatorname{tr}\left(g_\varepsilon^{-2} \sum_\alpha s_\alpha^* s_\alpha\right), \\ \langle\!\langle \Phi | A\Phi \rangle\!\rangle &= \sum_\alpha \operatorname{tr}\left(g_\varepsilon^{-1} s_\alpha^* \widetilde{F}(s_\alpha) g_\varepsilon\right) = \sum_\alpha \operatorname{tr}\left(s_\alpha^* \widetilde{F}(s_\alpha)\right) = f(S_{*,n}), \\ \langle\!\langle \Phi | A^2\Phi \rangle\!\rangle &= \sum_\alpha \left(g_\varepsilon \widetilde{F}(s_\alpha)^* \widetilde{F}(s_\alpha) g_\varepsilon\right) = \operatorname{tr}(g_\varepsilon^2 M), \\ \langle\!\langle \Phi | A^3\Phi \rangle\!\rangle &= \sum_\alpha \left(g_\varepsilon^2 \widetilde{F}(s_\alpha)^* \widetilde{F}(\widetilde{F}(s_\alpha) g_\varepsilon^2)\right) = f(S_{*,\varepsilon,n+1}), \end{aligned}$$

which implies monotonicity also in the Schrödinger picture. This completes the proof of the monotonicity of the subchannel power iteration in Theorem 3.2.9.

3.2.1.2 Fixed Points

We will now analyze if the subchannel power iteration does have fixed points when the objective value does not change. In principle, the iterated channel could still change in every iteration step, although the objective value does not increase. We will see that fixed points of the iteration exists, when the iteration results in a channel and not in a subchannel. Furthermore, we will extend the algorithm to ensure that this is indeed the case, so the iteration results in a channel after every step.

The case of equality of the objective in subsequent iterations provides two conditions. One is the equality in the last step of (3.54),

$$\operatorname{tr}\left(M^{1/2}\sum_{\alpha} s_\alpha^* s_\alpha\right) = \operatorname{tr} M^{1/2}, \tag{3.56}$$

with M given by equation (3.52). This means that the operator $\sum_\alpha s_\alpha^* s_\alpha$, which is the support projection of $S_n(\mathbb{1})$, acts like the identity on the support of M, or equivalently, on the support of $S_{n+1}(\mathbb{1})$,

$$\operatorname{supp} S_{n+1}(\mathbb{1}) \leq \operatorname{supp} S_n(\mathbb{1}). \tag{3.57}$$

Equality in the first inequality of (3.54) comes from near equality in Lemma 3.2.10. Equation (3.47) then leads us to,

$$\|A\phi - a\phi\| \to 0, \quad a = \frac{\langle\!\langle \phi | A\phi \rangle\!\rangle}{\langle\!\langle \phi | \phi \rangle\!\rangle}, \tag{3.58}$$

with an ε-dependence in A and ϕ. Recall that the Hilbert space is a direct sum over components labeled by α, so this limit statement can be written for each component separately,

$$\|h_\varepsilon \widetilde{F}(s_\alpha^*) - a h^{-1} s_\alpha^*\| \to 0. \tag{3.59}$$

If M has zero-eigenvalues, then we see from the defining equation (3.52) that h_ε is unbounded in the limit $\varepsilon \to 0$. So in general, we cannot conclude from (3.59) that

$$s^*_{\alpha,\varepsilon,n+1} = h_\varepsilon^2 \widetilde{F}(s_\alpha^*) \to a s^*_{\alpha,n}. \tag{3.60}$$

However, if we assume M to be invertible, i. e., the iterated S is always a channel and not a subchannel, equation (3.60) holds. For the eigenvalue, we then get

$$a = \frac{\langle\!\langle \phi | A\phi \rangle\!\rangle}{\langle\!\langle \phi | \phi \rangle\!\rangle} = \frac{f(S_n)}{\operatorname{tr}\left((M+\varepsilon\mathbb{1})^{1/2}\sum_\alpha s_\alpha^* s_\alpha\right)} \to \frac{f(S_n)}{\operatorname{tr}(M^{1/2})} = \frac{\sum_\alpha \operatorname{tr}\left(s_\alpha \widetilde{F}(s_\alpha^*)\right)}{\sum_\alpha \operatorname{tr}\left(\widetilde{F}(s_\alpha^*)^* \widetilde{F}(s_\alpha^*)\right)^{1/2}}.$$

where we used equation (3.56). So $a \geq 0$ and, since $f(S_n) = f(S_{n+1})$, we get $a = 0$ or $a = 1$. Thus, in the case of equality in (3.54), we have

$$s^*_{\alpha,n+1} = s^*_{\alpha,n}, \tag{3.61}$$

that is, we have $S_{n+1} = S_n$. So if M is invertible, we have indeed a fixed point in the subchannel power iteration.

On the other hand, if we found a vector Φ such that $M\Phi = 0$, $\|\Phi\| = 1$, that is, M is not invertible, we can add an additional Kraus operator

$$t = \sum_i a_i |\Psi_i\rangle\langle\Phi| \tag{3.62}$$

to the iterated channel S such that

$$S(\mathbb{1}) = t^*t + M^{-1/2}MM^{-1/2} \leq \mathbb{1}.$$

So S including the new Kraus operator is still a normalized subchannel. With this additional Kraus operator, the objective $f(S)$ has the additional term

$$\begin{aligned} \operatorname{tr}\left(t\widetilde{F}(t^*)\right) &= \operatorname{tr}\left(t\sum_{ij} F(|i\rangle\langle j|)t^*|j\rangle\langle i|\right) \\ &= \sum_{ijkl} a_k\bar{a}_l\langle\Phi|F(|i\rangle\langle j|)\Phi\rangle\langle\Psi_l|j\rangle\langle i|\Psi_k\rangle \\ &= \sum_{kl} a_k\bar{a}_l\langle\Phi|F(|\Psi_k\rangle\langle\Psi_l|)\Phi\rangle. \end{aligned} \tag{3.63}$$

This term does not increase the objective for any choice of Kraus operator of the form (3.62), as long as for all A

$$\langle\Phi|F(A)\Phi\rangle = 0. \tag{3.64}$$

We can rewrite this condition to a condition on the kernel of $F(\mathbb{1})$ with the following Lemma.

3.2.11 Lemma. *Given a completely positive and linear map $F\colon \mathcal{B}(\mathcal{H}_1) \to \mathcal{B}(\mathcal{H}_2)$ we have the equivalence*

$$\langle\Phi|F(A)\Phi\rangle = 0 \quad \forall A \in B(\mathcal{H}_1) \qquad \Leftrightarrow \qquad F(\mathbb{1})\Phi = 0$$

Proof. Let $\langle\Phi|F(A)\Phi\rangle = 0$ for all A. Then, in particular, we have $0 = \langle\Phi|F(\mathbb{1})\Phi\rangle = \|F(\mathbb{1})^{1/2}\Phi\|^2$, so $F(\mathbb{1})^{1/2}\Phi = 0$ and thus $F(\mathbb{1})\Phi = 0$.

Conversely, let $F(\mathbb{1})\Phi = 0$. Since $A \leq \|A\|\mathbb{1}$ for any positive operator A, and F maps positive operators to positive operators, we know that $F(A) \geq 0$ and that $F(\|A\|\mathbb{1} - A) \geq 0$. From linearity of F we obtain $\langle\Phi|F(A)\Phi\rangle \leq \|A\|\langle\Phi|F(\mathbb{1})\Phi\rangle$. Combined with the positivity of $F(A)$, we conclude that $\langle\Phi|F(A)\Phi\rangle = 0$ for all positive A. Furthermore, any operator A has a linear decomposition into positive operators [36], $A = \sum_i a_i A_i$, $A_i \geq 0$. Consequently, $F(A)\Phi = \sum_i c_i F(A_i)\Phi = 0$ for any operator A. ∎

This Lemma reduces the problem to the following: If $F(\mathbb{1})$ has a kernel and P is the projector onto this kernel, we can safely apply the iteration on the reduced Hilbert space $(1-P)\mathcal{H}_2$, since no vector in $P\mathcal{H}_2$, and thus no Kraus operator that maps from $P\mathcal{H}_2 \to \mathcal{H}_1$, can ever increase the objective value. But then, if we found a vector Φ in the iteration with the reduced Hilbert space such that $M\Phi = 0$, we can find an additional Kraus operator t of the form (3.62) that strictly increases the objective according to equation (3.63). If we add a Kraus operator for every

non-trivial vector from the kernel of M, we end up with an invertible normalization sum

$$M' = \sum_\alpha \widetilde{F}(s_\alpha^*)\widetilde{F}(s_\alpha^*)^* + \sum_\beta t_\beta^* t_\beta. \tag{3.65}$$

Furthermore, we can give a lower bound on the fidelity gain in advance. Suppose we have already reduced the Hilbert space $\mathcal{H}_2$ such that the kernel of $F(\mathbb{1})$ is trivial. Then, $F(\mathbb{1})$ has a smallest eigenvalue $\lambda_{\min}$ and

$$\langle\Phi|F(\mathbb{1})\Phi\rangle \geq \lambda_{\min}$$

for all Φ. We can rewrite this as average over a basis,

$$\frac{1}{d_1}\sum_i \langle\Phi|F(|i\rangle\langle i|)\Phi\rangle \geq \frac{\lambda_{\min}}{d_1},$$

where d_1 is the dimension of the Hilbert space $\mathcal{H}_1$. Since $\lambda_{\min}/d$ is an average value, we know there exists a vector $|i\rangle$ such that

$$\langle\Phi|F(|i\rangle\langle i|)\Phi\rangle \geq \frac{\lambda_{\min}}{d_1}.$$

Consequently, with the choice

$$t = |i\rangle\langle\Phi| \tag{3.66}$$

as additional Kraus operator, the gain in the objective (3.63) is lower bounded by $\lambda_{\min}/d_1$. Since the objective is bounded, this also implies that M cannot have a kernel if the objective value is above

$$\left(\max_S f(S)\right) - \lambda_{\min}/d_1.$$

Thus, M is invertible in this case and therefore (3.60) and (3.61) hold. From this it follows that a channel S, for which the global optimal objective value is attained, is a fixed point of the iteration. In particular, close to the global optimal objective value, the iterated channel has always full support on the reduced Hilbert space $(\mathbb{1} - P)\mathcal{H}_2$, leading to an iteration on channels rather than on subchannels. For this reason, we will call the modified subchannel power iteration the channel power iteration.

3.2.12 Definition (Channel Power Iteration)**.** Let $\mathcal{H}_1$, $\mathcal{H}_2$ be finite dimensional Hilbert spaces. Furthermore, let f be a linear positive objective on channels $S\colon \mathcal{B}(\mathcal{H}_1) \to \mathcal{B}(\mathcal{H}_2)$, given by the positive operator $\widetilde{F} \in \mathcal{B}(\mathcal{L}^2(\mathcal{H}_1, \mathcal{H}_2))$ according to equation (3.35). Then, the *channel power iteration* is given by:

1. Calculate the projection P onto the kernel of $F(\mathbb{1})$, where F is the given by (3.13). Calculate $\widetilde{F} \in \mathcal{B}(\mathcal{L}^2(\mathcal{H}_1, (\mathbb{1} - P)\mathcal{H}_2))$.

2. Start with a random channel S_0 with maximal number of linear independent Kraus operators.

3. Update the Kraus operators of S_n according to

$$\widetilde{S}'_{n+1} = \widetilde{F}\widetilde{S}_n\widetilde{F}.$$

4. If $M = S'_{n+1}(\mathbb{1}) = \sum_\alpha s'^*_{\alpha,n+1} s'_{\alpha,n+1}$ has a non-trivial kernel, add Kraus operators to S'_{n+1} according to (3.66) until $\ker M = \{0\}$.

5. Set $S_{n+1} = N(S'_{n+1})$, where N denotes the proper normalization depending on whether S is a channel in the Heisenberg picture or in the Schrödinger picture.

6. If $\|S_{n+1} - S_n\|_{\mathrm{cb}} > 0$ continue with step 3. Otherwise add Kraus operators to S_{n+1} according to equation (3.62), where the Φ are taken from the kernel of $F(\mathbb{1})$, until S_{n+1} becomes a channel $S_{n+1}\colon \mathcal{B}(\mathcal{H}_1) \to \mathcal{B}(\mathcal{H}_2)$.

Note that we already know from equation (3.63) that the objective value is neither increased nor decreased in the last step. If $\{|\beta\rangle\}$ is a basis of $P\mathcal{H}_2$, a choice of Kraus operators is given by $\{|\beta\rangle\langle\beta|\}$. Thus, the iteration always returns a proper normalized channel. Furthermore, close to the optimal value, the iteration is the same as the subchannel power iteration of Definition 3.2.8, since M has always full support, so no Kraus operators have to be added. A flowchart of the channel power iteration is shown in Figure 3.4.

3.2.13 Theorem (Channel Power Iteration)**.** *The channel power iteration given in Definition 3.2.12 has the following properties:*

1. *The iteration is monotone, that is, for successive iterates S_n, S_{n+1} we have $f(S_{n+1}) \geq f(S_n)$.*

2. *Every global optimal channel S is a fixed point of the iteration.*

3. *Let d_1 be the dimension of $\mathcal{H}_1$ and λ_{min} be the smallest eigenvalue of $F(\mathbb{1})$, where F is the completely positive map defined by equation* (3.13)*. Furthermore, let f_{max} be the global optimal value of f. Then, if*

$$f(S_n) > f_{max} - \lambda_{min}/d_1,$$

the operator $M = S'_{n+1}(\mathbb{1})$ in step 4 of the iteration is invertible, and therefore, the iteration does not add further Kraus operators (3.66) *to the channel.*

4. *The linearization of the iteration at a global optimal fixed point is stable, that is, the operator that acts on the perturbation as the linearization of the iteration is a contraction*[7]*.*

[7] A bounded operator is said to be a contraction if the operator norm of the operator is less than or equal to one.

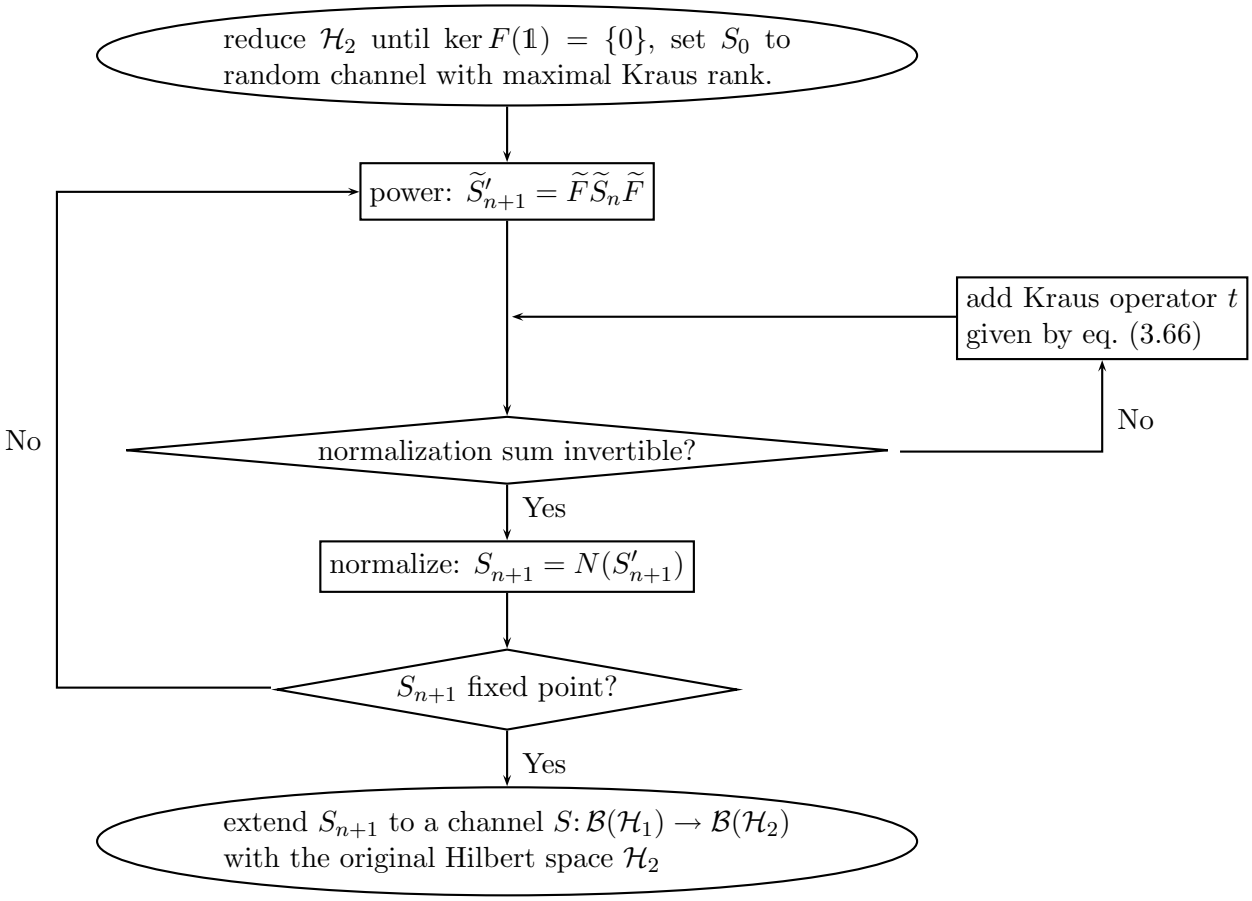

Figure 3.4: Flowchart of the channel power iteration from Definition 3.2.12.

Since the additional Kraus operators strictly increase the fidelity, the monotonicity of the channel power iteration, item 1 of the Theorem, follows from the monotonicity of the subchannel power iteration. Items 2 and 3 of the Theorem follow from the above discussion. The invertibility of the normalization sum near the global optimal value of the objective, which is guaranteed by item 3 of the Theorem, was required in the proof of the existence of fixed points. It will also become useful in the proof of the stability of the iteration, item 4 of the Theorem, which is presented below.

In summary, the channel power iteration is an iterative algorithm for the optimization of a linear, positive objective over the set of channels (Problem 3.2.2 on page 31). The iteration monotonically increases the objective value, and every global optimal channel is a stable fixed point of the iteration.

3.2.1.3 Stability

It remains to show that the iteration is stable close to the global optimum of the objective functional. Since the objective functional increases in every step and is bounded, we know that $f(S_n)$ converges. However, this does not ensure that the global optimum is found, even if the starting channel has an objective value that is very close to that optimum. So what we have to show is that for slight perturbations from the global optimal fixed point channel, the iteration reduces the perturbation, and hence brings the channel back to the global optimal fixed point.

To this end, we consider the variational problem of finding the maximum of (3.35). We will consider a fixed channel S and establish the conditions for it to be a local maximum of f. As for the monotonicity property, we regard a channel as tuple of Kraus operators. Let N be the number of Kraus operators of S. Then, the collection of Kraus operators $(s_1, \ldots, s_N)$ lies in the space of N-tuples of linear operators $s_\alpha : \mathcal{H}_1 \to \mathcal{H}_2$. We denote this space as Hilbert space $\mathcal{K}$, where the scalar product for two vectors $t, s \in \mathcal{K}$ is given by

$$\langle s|t\rangle = \sum_\alpha \operatorname{tr}(s_\alpha^* t_\alpha).$$

We will also consider $\mathcal{K}$ as a real Hilbert space $\mathcal{K}_\mathbb{R}$, with the scalar product $\langle s|t\rangle_\mathbb{R} = \operatorname{Re}\langle s|t\rangle$. A hermitian operator valued version of this scalar product is

$$\ll s, t\gg = \sum_\alpha (s_\alpha^* t_\alpha + t_\alpha^* s_\alpha). \tag{3.67}$$

For $t \in \mathcal{K}$ and $x \in \mathcal{B}(\mathcal{H}_1)$, we will write tx for the tuple $(tx)_\alpha = t_\alpha x$. For the variational analysis let $s(\eta)$ be a parametrized differentiable curve with $s(0) = s$,

$$s_\alpha(\eta) = s_\alpha + \eta\, \dot{s}_\alpha + \frac{1}{2}\eta^2\, \ddot{s}_\alpha + \mathcal{O}(\eta^3). \tag{3.68}$$

Note that complete positivity of the channel is implied by taking Kraus operators as variables rather than the channel itself. Thus, we no longer need to solve a variational problem with inequality constraint. The remaining constraint is

$$\sum_\alpha s_\alpha^* s_\alpha = \frac{1}{2} \prec s, s \succ = \mathbb{1}.$$

From this, we see that the channel constraint for $s(\eta)$ is satisfied to the first order, if and only if

$$\prec s, \dot{s} \succ = \sum_\alpha (s_\alpha^* \dot{s}_\alpha + \dot{s}_\alpha^* s_\alpha) = 0. \tag{3.69}$$

Note that this does not describe a $\mathbb{C}$-linear subspace, because both $\dot{s}_\alpha$ and $\dot{s}_\alpha^*$ appear in the condition. As an $\mathbb{R}$-linear subspace the set of admissible $\dot{s}_\alpha$ is described in the following Lemma.

3.2.14 Lemma. *Let $s \in \mathcal{K}$ be the tuple of Kraus operators of a channel S. Then the tangent space $\mathcal{T}$ at s to the manifold of channels is given by*

$$\mathcal{T} = \{ t \in \mathcal{K}_\mathbb{R} \,|\prec s, t \succ = 0 \}, \tag{3.70}$$

and its orthogonal complement in $\mathcal{K}_\mathbb{R}$ is

$$\mathcal{T}^\perp = \{ sx \in \mathcal{K} \,| x = x^* \in \mathcal{B}(\mathcal{H}_1) \}. \tag{3.71}$$

Moreover, the orthogonal decomposition $t = t^\parallel + t^\perp$ of a general $t \in \mathcal{K}_\mathbb{R}$ along this decomposition is given by

$$t_\alpha = \left(t_\alpha - \frac{1}{2} s_\alpha \prec s, t \succ \right) + \frac{1}{2} s_\alpha \prec s, t \succ . \tag{3.72}$$

Proof. The definition of the tangent space follows from (3.69). For the complement $\mathcal{T}^\perp$, we look at the scalar product of $t \in \mathcal{T}$ and $sx \in \mathcal{K}$, $x = x^* \in \mathcal{B}(\mathcal{H}_1)$ in $\mathcal{K}_\mathbb{R}$,

$$\begin{aligned} \langle t | sx \rangle_\mathbb{R} &= \mathrm{Re} \sum_\alpha \mathrm{tr}(t_\alpha^* s_\alpha x) = \frac{1}{2} \sum_\alpha \big(\mathrm{tr}(t_\alpha^* s_\alpha x) + \mathrm{tr}((s_\alpha x)^* t_\alpha) \big) \\ &= \frac{1}{2} \sum_\alpha \mathrm{tr} \left((t_\alpha^* s_\alpha + s_\alpha^* t_\alpha) x \right) = \frac{1}{2} \sum_\alpha \mathrm{tr} \left(\prec s, t \succ x \right) = 0. \end{aligned}$$

Thus, the orthogonal complement of the tangent space is indeed given by (3.71). The decomposition (3.72) follows from the fact that $\prec s, t \succ$ is hermitian, so $t^\perp$ is clearly of the form sx, $x = x^*$. Note that the calculation

$$\begin{aligned} \langle s | t^\parallel \rangle &= \sum_\alpha (s_\alpha^* t_\alpha^\parallel + t_\alpha^{\parallel *} s_\alpha) \\ &= \sum_\alpha \left(s_\alpha^* t_\alpha - \frac{1}{2} s_\alpha^* s_\alpha \prec s, t \succ + t_\alpha^* s_\alpha - \frac{1}{2} \prec s, t \succ s_\alpha^* s_\alpha \right) \\ &= \sum_\alpha \left(s_\alpha^* t_\alpha + t_\alpha^* s_\alpha \right) - \prec s, t \succ = 0 \end{aligned}$$

uses the normalization condition $\sum_\alpha s_\alpha^* s_\alpha = \mathbb{1}$. ∎

For the analysis of maxima, we also require that $s(\eta)$ satisfies the channel constraint to second order. For the η^2-term of the normalization sum $\sum_\alpha s_\alpha(\eta)^* s_\alpha(\eta)$ we get

$$0 = \sum_\alpha \Big(\dot{s}_\alpha^* \dot{s}_\alpha + \frac{1}{2} \left(s_\alpha^* \ddot{s}_\alpha + \ddot{s}_\alpha^* s_\alpha \right) \Big). \tag{3.73}$$

We will decompose the operator $\ddot{s}_\alpha$ as

$$\ddot{s}_\alpha = -s_\alpha \sum_\beta \dot{s}_\beta^* \dot{s}_\beta + r_\alpha, \tag{3.74}$$

with some $r \in \mathcal{K}$. Inserting this in (3.73) and using $\sum_\alpha s_\alpha^* s_\alpha = \mathbb{1}$ leads us to

$$\begin{aligned} 0 &= \sum_\alpha \Big(\dot{s}_\alpha^* \dot{s}_\alpha + \frac{1}{2} \big(s_\alpha^* (-s_\alpha \sum_\beta \dot{s}_\beta^* \dot{s}_\beta + r_\alpha) + (-\sum_\beta \dot{s}_\beta^* \dot{s}_\beta s_\alpha^* + r_\alpha^*) s_\alpha \big) \Big) \\ &= \sum_\alpha \dot{s}_\alpha^* \dot{s}_\alpha - \frac{1}{2} \sum_\beta \dot{s}_\beta^* \dot{s}_\beta + \frac{1}{2} \sum_\alpha s_\alpha^* r_\alpha - \frac{1}{2} \sum_\beta \dot{s}_\beta^* \dot{s}_\beta + \frac{1}{2} \sum_\alpha r_\alpha^* s_\alpha = \frac{1}{2} \prec s, r \succ, \end{aligned} \tag{3.75}$$

which is true if and only if $r \in \mathcal{T}$. So for $s(\eta)$ to be a channel to second order, we have the necessary conditions $\dot{s} \in \mathcal{T}$ and $\ddot{s}$ as given by equation (3.74) with $r \in \mathcal{T}$. On the other hand, these condition are also sufficient. Given $\dot{s}, r \in \mathcal{T}$, we can define

$$q_\alpha(\eta) = (s_\alpha + \eta \dot{s}_\alpha + \frac{1}{2} \eta^2 \ddot{s}_\alpha) h,$$

where

$$h = \Big(\sum_\alpha (s_\alpha + \eta \dot{s}_\alpha + \frac{1}{2} \eta^2 \ddot{s}_\alpha)^* (s_\alpha + \eta \dot{s}_\alpha + \frac{1}{2} \eta^2 \ddot{s}_\alpha) \Big)^{-1/2}$$

and $\ddot{s}_\alpha$ is given by equation (3.74). With this we have $\sum_\alpha q_\alpha^* q_\alpha = \mathbb{1}$, so q_α are indeed Kraus operators of a channel. Furthermore, since $h = (\mathbb{1} + \mathcal{O}(\eta^3))^{-1/2} = \mathbb{1} + \mathcal{O}(\eta^3)$, we know that $q_\alpha(\eta)$ has the decomposition (3.68). So every tuple

$$\boxed{(\dot{s}, r) \in \mathcal{T} \times \mathcal{T}}$$

defines a parametrized differentiable curve $s(\eta)$ of channels with $s(0) = s$ of the form (3.68) with $\ddot{s}_\alpha$ given by (3.74).

We now look at the optimality conditions for the first and second derivation of $f(S(\eta))$ at $\eta = 0$. For the first derivation, at $\eta = 0$, we get

$$\begin{aligned} \frac{df}{d\eta}\Big|_{\eta=0} &= \left[\frac{d}{d\eta} \sum_\alpha \operatorname{tr} \left(s_\alpha(\eta) \widetilde{F}(s_\alpha^*(\eta)) \right) \right]_{\eta=0} = \sum_\alpha \operatorname{tr} \left(\dot{s}_\alpha \widetilde{F}(s_\alpha^*) + s_\alpha \widetilde{F}(\dot{s}_\alpha^*) \right) \\ &= 2 \operatorname{Re} \sum_\alpha \operatorname{tr} \left(\dot{s}_\alpha \widetilde{F}(s_\alpha^*) \right) = 2 \langle \widetilde{F}(s^*)^* | \dot{s} \rangle_{\mathbb{R}}. \end{aligned} \tag{3.76}$$

Here we used that

$$\begin{aligned}\operatorname{tr}\left(s_\alpha \widetilde{F}(\dot{s}_\alpha^*)\right) &= \sum_{ij} \operatorname{tr}\left(s_\alpha F(|i\rangle\langle j|)\dot{s}_\alpha^*|j\rangle\langle i|\right)\\ &= \sum_{ij} \operatorname{tr}\left(|j\rangle\langle i|s_\alpha F(|i\rangle\langle j|)\dot{s}_\alpha^*\right)\\ &= \sum_{ij} \operatorname{tr}\left(\left(F(|j\rangle\langle i|)s_\alpha^*|i\rangle\langle j|\right)^*\dot{s}_\alpha^*\right) = \operatorname{tr}\left(\widetilde{F}(s_\alpha^*)^*\dot{s}_\alpha^*\right).\end{aligned}$$

As every $\dot{s} \in \mathcal{T}$ belongs to a differentiable curve of channels, (3.76) has to be zero for all $\dot{s} \in \mathcal{T}$ at an optimal value of f. So at an optimum we necessarily have

$$\boxed{\widetilde{F}(s^*)^* \in \mathcal{T}^\perp.} \tag{3.77}$$

For the second derivation, at $\eta = 0$, we get

$$\begin{aligned}\left.\frac{d^2 f}{d\eta^2}\right|_{\eta=0} &= \left[\frac{d^2}{d\eta^2}\sum_\alpha \operatorname{tr}\left(s_\alpha(\eta)\widetilde{F}(s_\alpha^*(\eta))\right)\right]_{\eta=0}\\ &= \sum_\alpha \operatorname{tr}\left(2\dot{s}_\alpha\widetilde{F}(\dot{s}_\alpha^*) + s_\alpha\widetilde{F}(\ddot{s}_\alpha^*) + \ddot{s}_\alpha\widetilde{F}(s_\alpha^*)\right)\\ &= 2\operatorname{Re}\sum_\alpha \operatorname{tr}\left(\dot{s}_\alpha\widetilde{F}(\dot{s}_\alpha^*) + \ddot{s}_\alpha\widetilde{F}(s_\alpha^*)\right).\end{aligned} \tag{3.78}$$

We will now insert (3.74) into this equation. For a stationary point we can drop the term proportional to r since

$$\operatorname{Re}\sum_\alpha \operatorname{tr}\left(r_\alpha\widetilde{F}(s_\alpha^*)\right) = \langle\widetilde{F}(s^*)^*|r\rangle_{\mathbb{R}} = 0$$

from the condition (3.77) from the first derivation. Thus we obtain

$$\left.\frac{d^2 f}{d\eta^2}\right|_{\eta=0} = 2\operatorname{Re}\sum_\alpha \operatorname{tr}\left(\dot{s}_\alpha\widetilde{F}(\dot{s}_\alpha^*) - \left(\sum_\beta \dot{s}_\beta^*\dot{s}_\beta\right)\widetilde{F}(s_\alpha^*)s_\alpha\right), \tag{3.79}$$

where we used the cyclicity of the trace. As $\sum_\beta \dot{s}_\beta^*\dot{s}_\beta$ is hermitian, the sum depends only on the hermitian part $\frac{1}{2}\lessdot s, \widetilde{F}(s^*)^*\gtrdot$ of $\sum_\alpha \widetilde{F}(s_\alpha^*)s_\alpha$. Equation (3.79) then reads

$$\left.\frac{d^2 f}{d\eta^2}\right|_{\eta=0} = 2\operatorname{Re}\sum_\alpha \operatorname{tr}\left(\dot{s}_\alpha\widetilde{F}(\dot{s}_\alpha^*) - \frac{1}{2}\dot{s}_\alpha^*\dot{s}_\alpha \lessdot s, \widetilde{F}(s^*)^*\gtrdot\right), \tag{3.80}$$

where the summation index β was renamed to α. At a maximum, these derivatives have to be negative for all $\dot{s} \in \mathcal{T}$, i. e., s corresponds to a local maximum, if

$$\boxed{\sum_\alpha \operatorname{tr}\left(\dot{s}_\alpha\widetilde{F}(\dot{s}_\alpha^*)\right) \leq \frac{1}{2}\sum_\alpha \operatorname{tr}\left(\dot{s}_\alpha^*\dot{s}_\alpha \lessdot s, \widetilde{F}(s^*)^* \gtrdot\right).} \tag{3.81}$$

So the optimality conditions are given by equation (3.77) and (3.81). As the objective is linear in the set of channels, it has a single global optimum. For any two channels S and T for which $f(S) > f(T)$, the convex combination $(\eta S + (1-\eta)T)$ corresponds to a curve of channels along which the objective functional strictly increases with η, so $f(T)$ cannot be a local optimum. However, from the largest eigenvalue iteration we know that there may be fixed points for which the global maximum is not reached, namely, when the starting vector has no component in direction of the eigenvector of the largest eigenvalue. But a small perturbation of such a fixed point is sufficient to get the iteration going again, and therefore generically going to the global optimum. In our case, this means that we have to start the iteration with a channel that has full Kraus rank.

We will now look at the linearization of the iteration at the global optimum of the objective functional. As we already proofed items 2 and 3 of Theorem 3.2.13, we know that the operator $M = S'_{n+1}(\mathbb{1})$ is invertible near the optimum and that each optimal channel is a fixed point of the iteration. In particular, the iteration does not add further Kraus operators near the optimum, the channel power iteration is the same as the subchannel power iteration. So let $s \in \mathcal{K}$ belong to a global optimal channel, then equation (3.61) holds, or equivalently,

$$\widetilde{F}(s^*_\alpha) = s^*_\alpha. \tag{3.82}$$

We look at perturbations $s(\eta)$ of this channel of the form (3.68), $s_\alpha(\eta) = s_\alpha + \eta \dot{s}_\alpha + \mathcal{O}(\eta^2)$. Each step of the iteration consists of an application of the operator $\widetilde{F}$ and a normalization operation, so the iteration step becomes

$$s^*_{\alpha,n+1} = M^{-1/2}\widetilde{F}(s^*_{\alpha,n})$$

near the global optimum.

The first part of the iteration step is the application of the already linear operator $\widetilde{F}$, which leads us to

$$\widetilde{F}\left(s^*_\alpha(\eta)\right) = s^*_\alpha + \eta\widetilde{F}(\dot{s}^*_\alpha) + \mathcal{O}(\eta^2).$$

Here we used the fixed point property (3.82) and the linearity of $\widetilde{F}$.

Next, we have to compute the operator M,

$$\begin{aligned} M(\eta) &= \sum_\alpha \widetilde{F}\left(s^*_\alpha(\eta)\right)\widetilde{F}\left(s^*_\alpha(\eta)\right)^* \\ &= \mathbb{1} + \eta\sum_\alpha \left(s^*_\alpha \widetilde{F}(\dot{s}^*_\alpha)^* + \widetilde{F}(\dot{s}^*_\alpha)s_\alpha\right) + \mathcal{O}(\eta^2) \\ &= \mathbb{1} + \eta \prec s, \widetilde{F}(\dot{s}^*)^* \succ + \mathcal{O}(\eta^2). \end{aligned}$$

To compute the inverse of the square root to the first order, we make the ansatz

$M(\eta)^{-1/2} = \mathbb{1} + \eta W + \mathcal{O}(\eta^2)$ and solve the equation

$$\begin{aligned}\mathbb{1} + \mathcal{O}(\eta^2) &= \left(M(\eta)^{-1/2}\right)^2 M(\eta) = (\mathbb{1} + \eta W)(\mathbb{1} + \eta W)M(\eta) \\ &= (\mathbb{1} + 2\eta W)M(\eta) + \mathcal{O}(\eta^2) = \mathbb{1} + \eta\left(\prec s, \widetilde{F}(\dot{s}^*)^* \succ + 2W\right) + \mathcal{O}(\eta^2).\end{aligned}$$

This leads us to $W = -\frac{1}{2} \prec s, \widetilde{F}(\dot{s}^*)^* \succ$, and hence

$$M(\eta)^{1/2} = \mathbb{1} - \eta\frac{1}{2} \prec s, \widetilde{F}(\dot{s}^*)^* \succ + \mathcal{O}(\eta^2).$$

Finally, we have to apply $M(\eta)^{-1/2}$ to $\widetilde{F}(s(\eta)^*)$, so the linearization of a complete iteration step becomes

$$\begin{aligned}M(\eta)^{-1/2}\widetilde{F}(s(\eta)^*) &= \left(\mathbb{1} - \eta\frac{1}{2} \prec s, \widetilde{F}(\dot{s}^*)^* \succ\right)\left(s^*_\alpha + \eta\widetilde{F}(\dot{s}^*_\alpha)\right) + \mathcal{O}(\eta^2) \\ &= s^*_\alpha + \eta\left(\widetilde{F}(\dot{s}^*_\alpha) - \frac{1}{2} \prec s, \widetilde{F}(\dot{s}^*)^* \succ s^*_\alpha\right) + \mathcal{O}(\eta^2).\end{aligned}$$

If we compare this result with Lemma 3.2.14, we see that the linear term corresponds to the projection of $\widetilde{F}(\dot{s}^*_\alpha)^*$ onto the tangent space $\mathcal{T}$. Thus, that term can be written in the form $PB\dot{s}$ with a projection P and an operator B given by $(B\dot{s})_\alpha = \widetilde{F}(\dot{s}^*_\alpha)$. With this definition, the linearization of the iteration acts on the perturbation as

$$\dot{s}^*_{n+1} = PB\dot{s}^*_n. \tag{3.83}$$

Since the initial perturbation $\dot{s}$ will be in the tangent space, we only have to consider the powers of the operators PBP for the analysis of the asymptotic behavior of the iteration near a global optimal fixed point. If PBP is a contraction, that is, $\|PBP\| \leq 1$, the iteration is stable, and even asymptotically stable if $\|PBP\| < 1$. Since PBP is positive, all eigenvalues are non-negative, so no oscillatory behavior is possible. On the other hand, if $\|PBP\| > 1$, there would be an initial deviation from s which blows up exponentially. However, since

$$\frac{1}{2} \prec s, \widetilde{F}(s^*)^* \succ = \frac{1}{2}\sum_\alpha \left(s^*_\alpha \widetilde{F}(s^*_\alpha)^* + \widetilde{F}(s^*_\alpha)s_\alpha\right) = \frac{1}{2}\sum_\alpha (2s^*_\alpha s_\alpha) = \mathbb{1}$$

at a fixed point s, the optimality condition (3.81) becomes

$$\sum_\alpha \operatorname{tr}\left(\dot{s}_\alpha \widetilde{F}(\dot{s}^*_\alpha)\right) \leq \sum_\alpha \operatorname{tr}\left(\dot{s}^*_\alpha \dot{s}_\alpha\right)$$

and we obtain

$$\langle \dot{s}|PBP|\dot{s}\rangle \leq \langle \dot{s}|\dot{s}\rangle,$$

which, as PBP is positive, is the same as $\|PBP\| \leq 1$. Thus the iteration is stable.

In summary, the linearization of the iteration applied to a deviation from a global optimal fixed point becomes the application of the linear operator $\widetilde{F}$ followed by a

projection to the tangent space of the channel manifold. The optimality condition for the fixed point implies that the overall operation is a contraction, and thus, that the iteration is stable. This shows item 4 of Theorem 3.2.13 and concludes the proof of Theorem 3.2.13 and Theorem 3.2.9.

3.2.1.4 Implementation

We will now estimate the computational cost of a single iteration step of the subchannel power iteration based on the number of floating point operations (flops). For this, we treat operations on complex numbers as a single flop. Note that flop counts give only a rough estimate on the runtime on modern (classical) computer hardware, as e. g., possible parallelism, cachability, and locality of reference can all have dramatic impact on the computation time.

For a channel $S\colon \mathcal{B}(\mathcal{H}_1) \to \mathcal{B}(\mathcal{H}_2)$, the subchannel power iteration requires the channel S as well as the Jamiolkowsky dual $\widetilde{S}$ as data. Therefore we choose, according to equation (3.14),

$$s(a,i,b,j) = \langle a|S(|i\rangle\langle j|)|b\rangle = \langle a|\widetilde{S}(|b\rangle\langle j|)|i\rangle \tag{3.84}$$

as data type to store the channel S. This requires $(d_1 d_2)^2$ storage units, where $d_1 = \dim \mathcal{H}_1$ and $d_2 = \dim \mathcal{H}_2$. Note that by taking the elements columnwise from s, we can reshape[8] s and obtain the $(d_1 d_2) \times (d_1 d_2)$-matrix representation $\widehat{S}$ defined in (3.15). We will make no assumptions about the structure of this matrix (e. g., sparsity, rank).

For the flop-count analysis, we will divide the subchannel power iteration of Definition 3.2.8 into three steps:

1. Apply the operator $\widetilde{F}$, $\widetilde{S} \mapsto \widetilde{F}\widetilde{S}\widetilde{F}$.
2. Compute the pseudo-inverse $S(\mathbb{1})^{-1/2}$.
3. Apply the pseudo-inverse to normalize the new iterate,

$$S(x) \mapsto S(\mathbb{1})^{-1/2} S(x) S(\mathbb{1})^{-1/2}.$$

The first step is equivalent to a matrix multiplication with the reshaped s. For the matrix multiplication of $n \times n$-matrices A and B, we have to compute

$$\sum_k \langle i|A|k\rangle\langle k|B|j\rangle \tag{3.85}$$

for every pair (i,j). As the computation of (3.85) requires n multiplications and $(n-1)$ additions, we have a total flop count of $n^2(2n-1)$. In our estimation, we

[8] This corresponds to the action of the Matlab command *reshape(s, [$d_1 d_2$, $d_1 d_2$])*.

are only interested in the order of the flop count, so we treat unconstrained matrix multiplication as $\mathcal{O}(n^3)$. Thus, the first step has a flop count of $\mathcal{O}((d_1 d_2)^3)$.

The second step involves the computation of the pseudo-inverse of the square root of $S(\mathbb{1})$. The computation of $\langle a|S(\mathbb{1})|b\rangle = \sum_i \langle a|S(|i\rangle\langle i|)|b\rangle$ needs $d_1 d_2^2$ flops. For the computation of $S(\mathbb{1})^{-1/2}$, we need the eigenvalue decomposition of $S(\mathbb{1})$. We can exploit the fact that $S(\mathbb{1}) \geq 0$. Thus, we can convert the problem of finding the eigenvalue decomposition of $S(\mathbb{1})$ into the problem of finding the eigenvalue decomposition of the real symmetric matrix

$$\begin{pmatrix} A & -B \\ B & A \end{pmatrix},$$

where $S(\mathbb{1}) = A + iB$. As this increases the complexity only by a factor [68], the order of floating point operations is still appropriately estimated. Several algorithms for computing the eigenvalue decomposition for real symmetric matrices exist [68]. We will use the Jacobi method (Contribution II/1 by H. Rutishauser, [68]) for the flop count, although, as noted by Rutishauser, a combination of Householder-transformation together with methods for tridiagonal matrices is more efficient for large matrices. The Jacobi algorithm has quadratic convergence in the number of Jacobi rotations. Jacobi rotations are of the form

$$A \mapsto U^T A U,$$

where $U = U(p, q, \phi)$ is an orthogonal matrix which deviates form the unit matrix only in the elements $\langle p|Up\rangle = \langle q|Uq\rangle = \cos(\phi)$ and $\langle p|Uq\rangle = -\langle q|Up\rangle = \sin(\phi)$. Thus we have $\mathcal{O}(d_2^2)$ steps and $\mathcal{O}(d_2)$ operations per step, which leads to a flop count of $\mathcal{O}(d_2^3)$ for the eigenvalue decomposition. Given the eigenvalues λ_i, we compute the eigenvalues of $S(\mathbb{1})^{1/2}$ as $1/\sqrt{\lambda_i}$, if $\lambda_i > 0$. Otherwise, we will add further Kraus operators in the channel power iteration (Def. 3.2.12), or set the new eigenvalues to zero in the subchannel power iteration (Def. 3.2.8). As the gain in the objective for each additional Kraus operator in the channel power iteration is lower bounded, $\lambda = 0$ will not happen close to the optimal channel. Therefore, we will not consider the flops needed for finding additional Kraus operators. Finally, given the eigenvalues of $S(\mathbb{1})^{1/2}$ as diagonal matrix D, we have to compute $S(\mathbb{1})^{1/2} = VDV^T$, where V is the product of the Jacobi rotation matrices U, that can already be computed along with the Jacobi rotation itself. As matrix multiplication, the last step has also at most $\mathcal{O}(d_2^3)$ flops.

In the third step, we compute $\langle a|S(\mathbb{1})^{-1/2} S(|i\rangle\langle j|) S(\mathbb{1})^{-1/2}|b\rangle$ for all pairs (a, i, b, j) with the two equations,

$$\langle a|S(|i\rangle\langle j|)S(\mathbb{1})^{-1/2}|b\rangle = \sum_{k=1}^{d_2} \langle a|S(|i\rangle\langle j|)|k\rangle\langle k|S(\mathbb{1})^{-1/2}|b\rangle$$

and

$$\langle a|S(\mathbb{1})^{-1/2}S(|i\rangle\langle j|)S(\mathbb{1})^{-1/2}|b\rangle = \sum_{l=1}^{d_2}\langle a|S(\mathbb{1})^{-1/2}|l\rangle\langle l|S(|i\rangle\langle j|)S(\mathbb{1})^{-1/2}|b\rangle.$$

The above equations have both the same flop count, so the third step requires $\mathcal{O}(d_1^2 d_2^3)$ floating point operations. Note that if we would solve this with a single equation, we would unnecessarily compute the second sum due, to the decomposition of the unity, for every term of the first sum. This would give an additional factor d_2.

In total, we have that every iteration step costs

$$\mathcal{O}((d_1 d_2)^3) + \mathcal{O}(d_2^3) + \mathcal{O}(d_1^2 d_2^3) = \mathcal{O}((d_1 d_2)^3) \tag{3.86}$$

floating point operations. From the comparison with the power iteration, we expect linear convergence of the overall iteration depending on the ratio of the two largest eigenvalues of $\widetilde{F}$. However, as we haven't developed an explicit lower bound on the rate of convergence, we cannot give a flop count estimate for the total iteration.

The implementation of the seesaw and the (sub)channel power iterations are straight forward. The channel is stored using the array (3.84). This has the advantage that calculations with the Jamiolkowsky dual for the power step and calculations with the usual channel representation in the normalization step can both be computed efficiently. For this, a new class `HSChannel`[9] was added to the existing object oriented framework of my diploma thesis [53]. The class and its connection to the existing channel representations with Stinespring isometry or Kraus operators are shown in the UML diagram in Figure 3.5. For performance reason, the class stores the input and output dimension of the channel in addition to the channel array (3.84). From equation (3.16), $\widetilde{T} = \sum_\alpha \big|t_\alpha^*\big\rangle\!\big\rangle\big\langle\!\big\langle t_\alpha^*\big|$, we see that every Kraus decomposition belongs to a decomposition $\widehat{T} = BB^*$, where the columns of B are given by the Kraus operators t_α^*. The converse is also true [65], in particular, the eigendecomposition, where we combine the eigenvalues and eigenvectors to obtain the form $\widehat{T} = BB^*$, belongs to Kraus operators that are linear independent as they are orthogonal in the Hilbert-Schmidt scalar product $\big\langle\!\big\langle x\big|y\big\rangle\!\big\rangle = \operatorname{tr} x^* y$. This is interesting, as we can deduce from the number of independent Kraus operators whether a channel is isometric or homomorphic. The matrix $\widehat{T}$ is obtained from the array (3.84) via a simple `reshape` operation. Thus, `HSChannel` implements the `KrausChannel` interface with $\mathcal{O}((d_1 d_2)^3)$ flops for the eigenvalue decomposition. Furthermore, `HSChannel` has a method `dual` that maps Schrödinger channels to Heisenberg channels and vice versa, i. e., $t_\alpha \mapsto t_\alpha^*$, via the equation

$$t(a,i,b,j) = \langle a|T(|i\rangle\langle j|)|b\rangle = \langle j|T_*(|b\rangle\langle a|)|i\rangle = t_*(j,b,i,a).$$

[9] The name was chosen to indicate that $\widetilde{T}$ operates on $\mathcal{L}^2(\mathcal{H}_1, \mathcal{H}_2)$.

Figure 3.5: UML class diagram of channel classes. Updated version of the diagram in [53]. The Channel interface has a new operation `addChannel` for convex combinations of channels. The new class `HSChannel` holds the matrix (3.84) for easy calculations in Jamiolkowsky dual and standard channel representation alike.

The subchannel power iteration (Definition 3.2.8 on page 38) was implemented for the different channel types via the classes `HSDecoderIteration`, `HSEncoderIteration`, and `HSNoiseIteration`. The additional steps of the channel power iteration (Definition 3.2.12) were omitted, as in practice, the iteration results turned out to be always a channel and never a subchannel. From equations (3.28) and (3.30) we see that the iterated channel D in the decoder iteration is in the Heisenberg picture, while the iterated channel E_* in the encoder iteration is in the Schrödinger picture. That is, the encoder and decoder iterations have different normalizations. Instead of normalizing E_* directly, the dual E was normalized, so the encoder iteration only differs from the decoder iteration in two additional calls to `dual`, before and after the normalization step. Therefore, we restrict the discussion to the decoder iteration.

The class `HSDecoderIteration` takes the concatenated channel ET as well as the starting channel D as arguments. In compliance with equation (3.28), it stores the dual $(ET)_*$ as map F, and D as S, the channel to optimize. The class provides the method `iterate`, that iterates until the gain in fidelity is below a given threshold. Each step of the iteration is given by the following method, the implementation of the subchannel power iteration step of Definition 3.2.8.

```
function this = iterationStep( this )
% computes one step of the iteration

% S^ -> F^*S^*F^
in = getInputdim(this.hsS);
out = getOutputdim(this.hsS);
sHsMap = reshape(getHilbertSchmidtMap(this.hsS),[in*out,in*out]);
fHsMap = reshape(getHilbertSchmidtMap(this.hsF),[in*out,in*out]);

newSHsMap = reshape(fHsMap*sHsMap*fHsMap,[out,in,out,in]);

% normalize D
% S(x) |-> S(1)^(-1/2)S(x)S(1)^(-1/2)

% S(1) = sum_i S(|i><i|)
newHsDOfOne = zeros(out,out);
for j=1:in
  newHsDOfOne = newHsDOfOne + squeeze(newSHsMap(:,j,:,j));
end

% S(1)^(-1/2)
n = pinv(sqrtm(newHsDOfOne));

% S(|k><j|) -> S(1)^(-1/2)S(|k><j|)S(1)^(-1/2)
ndn = zeros(out,in,out,in);
for k=1:in
    for j=1:in
        ndn(:,k,:,j) = n*squeeze(newSHsMap(:,k,:,j))*n;
```

```
    end
end

this.hsS = CHSChannel(ndn);
```

The reshape operations are used to turn the array (3.84) into the Jamiolkowsky matrix and vice versa. Matlab's internal functions are used for the computation of the pseudo inverse $D(\mathbb{1})^{-1/2}$. As the code dimension, and therefore the input dimension of the decoder in Heisenberg picture, is usually small compared to the noise dimension, e. g., `in`$= 2$ in the qubit case, there would be little benefit from further code vectorization[10]. However, readability of the code would be greatly reduced.

Finally, the seesaw iteration takes the noisy channel, the dimension of the code space, and optionally the desired accuracy, and returns the channel fidelity as well as the code (E, D) found.

```
function [fidelity, encoderChannel, decoderChannel] = ...
  seesawPower(inDim, noiseChannel, seesawAccuracy)

% accuracy
if (nargin == 2)
  seesawAccuracy = 10^(-4);
end

encoderChannel = CHSChannel( ...
                   CRandomChannel(getOutputdim(noiseChannel),inDim));
decoderChannel = CHSChannel( ...
                   CRandomChannel(inDim,getInputdim(noiseChannel)));

endIteration = false;
while not(endIteration);

  ET = concatenate(encoderChannel, noiseChannel);
  decoderIteration = CHSDecoderIteration(ET, decoderChannel);
  decoderIteration = iterate(decoderIteration);
  decoderObjectiveValue = computeObjectiveValue(decoderIteration);
  decoderChannel = getDecoder(decoderIteration);

  TD = concatenate(noiseChannel, decoderChannel);
  encoderIteration = CHSEncoderIteration(TD, encoderChannel);
  encoderIteration = iterate(encoderIteration);
  encoderObjectiveValue = computeObjectiveValue(encoderIteration);
  encoderChannel = getEncoder(encoderIteration);

  if(abs(encoderObjectiveValue - decoderObjectiveValue) < seesawAccuracy)
    endIteration = true;
```

[10] This can also be seen using Matlab's profiler.

```
  end

end
fidelity = decoderObjectiveValue/inDim^2;
```

3.2.1.5 Conclusion

The optimization of a linear functional over the set of channels can be solved via a modified version of the power iteration, the channel power iteration or subchannel power iteration. Both algorithms improve the linear functional in every step and global optimal solutions are stable fixed points of the channel power iteration. The iterations do not require any special properties of the channels like symmetry. Furthermore, it is possible to limit the number of independent Kraus operators of the solution. Both methods have a high numerical stability as errors do not accumulate over iteration steps. The subchannel power iteration has an easy implementation and a fast runtime of the iteration step compared to the semidefinite programming approach about to be mentioned.

The iterations can be used to find a local optimal quantum error correcting code for a given noisy channel by alternately optimizing the encoding and decoding channel. As the iterations do not require any special properties of the channels, they can find codes that do not rely on the Knill-Laflamme condition (3.2). Furthermore, one can restrict the encoding to be isometric due to the ability to limit the number of independent Kraus operators. This option is particularly interesting for larger systems as the problem size grows exponentially in the number of qubits.

3.2.2 Semidefinite Programming

As already mentioned above, the single step in the seesaw iteration, that is, the optimization of a linear objective over the set of channels (Problem 3.2.2), can also be solved using semidefinite programming. This fact was first published by Audenaert and De Moor [63] and was recently rediscovered by Fletcher, Shor, and Win [69] and Kosut and Lidar [70]. The main virtue of using semidefinite programming is that, under certain technical conditions, the numerical result is guaranteed to lie in an arbitrarily small interval around the true value of the global optimum.

We already identified the structure of the semidefinite constraint for subchannels $T\colon \mathcal{B}(\mathcal{H}_1) \to \mathcal{B}(\mathcal{H}_2)$ in the Heisenberg picture in equation (3.21) on page 33,

$$\widehat{T} \oplus (\mathbb{1} - \mathrm{tr}_{\mathcal{H}_1} \widehat{T}) \geq 0.$$

Here $\widehat{T}$ is the matrix representation of the Jamiolkowsky dual $\widetilde{T}$ of T, as defined in equation (3.15). The constraint ensures that T is a completely positive map and

that $T(\mathbb{1}) \leq \mathbb{1}$. Hence, T is a subchannel in the Heisenberg picture. With another direct summand,

$$\widehat{T} \oplus (\mathbb{1} - \mathrm{tr}_{\mathcal{H}_1} \widehat{T}) \oplus (-\mathbb{1} + \mathrm{tr}_{\mathcal{H}_1} \widehat{T}) \geq 0,$$

we can even restrict to the equality constraint $T(\mathbb{1}) = \mathbb{1}$. It is more natural to write this in the standard primal form of a conic program (see p. 266 in [39]),

$$\min \{ \langle c|x\rangle \,|Ax = b,\, x \in \mathcal{C} \}, \tag{3.87}$$

where $\langle c|x\rangle$ is a linear functional, $Ax = b$ is a linear system of equations, and $\mathcal{C}$ is a pointed convex cone. In our case, we have $x = \widehat{T}$, $\langle c|x\rangle = -\mathrm{tr}(\widehat{F}\widehat{T})$ with $\widehat{F}$ defined by equation (3.22), $Ax = \mathrm{tr}_{\mathcal{H}_1} \widehat{T}$, $b = \mathbb{1}$, and $\mathcal{C}$ is the cone of positive semidefinite matrices $\widehat{T} \geq 0$. Usual implementations require x to be in vector form, so we take x to be the vector that results from stacking the columns of the corresponding matrix[11]. For channels in the Schrödinger picture $T_*\colon \mathcal{B}(\mathcal{H}_1) \to \mathcal{B}(\mathcal{H}_2)$, the only difference is that $Ax = \mathrm{tr}_{\mathcal{H}_2} \widehat{T_*}$ as we see from Lemma 3.2.5.

So, Problem 3.2.2 has the form of the semidefinite program

$$\begin{aligned} \text{minimize} \quad & -f(T) = -\mathrm{tr}\left(\widehat{F}\widehat{T}\right) \\ \text{subject to} \quad & \mathrm{tr}_{\mathcal{H}_1} \widehat{T} = \mathbb{1}, \\ & \widehat{T} \geq 0. \end{aligned} \tag{3.88}$$

The corresponding dual problem is

$$\max \{ \langle b|y\rangle \,|c - A^* y \in \mathcal{C}^* \}, \tag{3.89}$$

where $\mathcal{C}^*$ denotes the dual cone of $\mathcal{C}$ defined by $\mathcal{C}^* = \{y \,|\langle x|y\rangle \geq 0 \forall x \in \mathcal{C}\}$. Here $\mathcal{C} = \mathcal{C}^*$, since the cone of positive semidefinite matrices is self-dual. Audenaert and De Moor [63] wrote (3.88) in the standard semidefinite form

$$\min \left\{ \langle c|x\rangle \,\middle|\, F(x) = F_0 + \sum_i x_i F_i \geq 0 \right\}, \tag{3.90}$$

with the dual problem

$$\max \{ -\mathrm{tr}(F_0 Z) \,|Z \geq 0; \mathrm{tr}(F_i Z) = c_i \}. \tag{3.91}$$

They made the partial trace constraint implicit via their choice of parametrization and showed that $F_0 = \mathbb{1}_{\mathcal{H}_1 \otimes \mathcal{H}_2} / \dim \mathcal{H}_2$ for channels in the Schrödinger picture. From this we immediately see that $F(0) > 0$, that is, the primal problem is strictly feasible for $x = 0$. This implies (Theorem 3.1 [33]) that the duality gap[12] is zero, i. e., the optimal values of the primal and dual objectives are equal. As the dual

[11] Column stacking is also known as vectorization.

[12] The duality gap is the difference between the primal and dual objective

objective is always less than or equal to the primal objective, each feasible[13] pair (x, Z) results in an interval $[-\operatorname{tr}(F_0 Z), \langle c|x\rangle]$ in which the true optimal value lies. So each solution Z for the dual problem for which the duality gap is $\langle c|x\rangle + \operatorname{tr}(F_0 Z) \leq \varepsilon$ certifies that x is optimal up to ε.

To estimate the runtime of the primal problem, we use the floating point operations (flop) count of the semidefinite programming implementation given in chapter 11.8.3 of [39]. In the primal form (3.90) with $x \in \mathbb{R}^n$ and complex hermitian $p \times p$-matrices F_i, the dominating order in the flop count for a single iteration step is

$$\max\{np^3, n^2p^2, n^3\}. \tag{3.92}$$

The Jamiolkowsky dual $\widetilde{T}$, and therefore x, has about $(d_1 d_2)^2$ real parameters, where $d_1 = \dim \mathcal{H}_1$ and $d_2 = \dim \mathcal{H}_2$, so we have $n = \mathcal{O}((d_1 d_2)^2)$. Observe that we can also parametrize $\widetilde{F}$ using $d_1 d_2$ real parameters for the diagonal and $(d_1 d_2/2(d_1 d_2 + 1) - d_1 d_2)$ complex parameters for the off-diagonal matrix elements. This parametrization would be more comparable to the power iteration, however, the dominant order would still be $(d_1 d_2)^2$. As the parametrization by Audenaert and De Moor implicitly guarantees the correct partial trace, the semidefinite constraint is only $\widetilde{T} \geq 0$, and therefore we have $p = d_1 d_2$. In summary, we obtain a flop count of

$$\mathcal{O}((d_1 d_2)^6)$$

for an iteration step of the primal problem. This is more expensive than the $\mathcal{O}((d_1 d_2)^3)$ from (3.86) on page 56 for the power iteration.

As shown in [63], the dual problem can be reduced to a semidefinite program with $p = d_1 d_2$ and only $n = d_1^2$ unknown parameters instead of $(d_1 d_2)^2$ parameters for Z in (3.91). From (3.92) we obtain a flop count with dominant order

$$\max\{d_1^2 (d_1 d_2)^3, d_1^4 (d_1 d_2)^2\}$$

for a single iteration step of the dual problem. Although this is still worse than the power iteration, it is significantly better than the computational cost of the primal problem.

The implementation was done using SeDuMi [47]. SeDuMi solves conic problems in the form (3.87), which is why we formulated the optimization as conic problem in the first place. The code directly implements (3.88).

```
function [fidelity, encoderChannel, decoderChannel] = ...
  seesawSDP(inDim, noiseChannel, seesawAccuracy, sdpAccuracy)
% Finds the optimal Encoder–Decoder pair with semidefinite programming.
%
```

[13] That is, (x, Z) satisfies the constraints of the primal and dual problem, respectively.

```
% Optimizes the channel fidelity with a seesaw iteration, using
% semidefinite programming for the single optimizations. The seesaw
% iteration is started with a random encoder and decoder channel.
% Alternately the encoder and decoder are fixed and the decoder and
% encoder are optimized, respectively.
%
% seesawSDP(inputDimension, noiseChannel, seesawAccuracy, sdpAccuracy)
%
% seesawSDP(inputDimension, noiseChannel, seesawAccuracy)
% Single optimization of encoder and decoder ends, if the fidelity gain
% is below seesawAccuracy/10
%
% seesawSDP(inputDimension, noiseChannel)
% Seesaw iteration ends, if the fidelity gain is below 1e-4. Single
% optimization of encoder and decoder ends,if the fidelity gain is below
% 1e-5.
%
% inDim : dimension of the Hilbert space to code into the channel.
% noiseChannel : channel in Heisenberg picture to use by the coding.
% seesawAccuracy : seesaw iteration ends, if the fidelity gain is below
%   this value.
% sdpAccuracy : single optimization of encoder and decoder ends, if the
%   fidelity gain is below this value.
% return : fidelity of the optimal solution, encoder and decoder channel
%   in Heisenberg picture.

% accuracy
if (nargin == 2)
  seesawAccuracy = 10^(-4);
  sdpAccuracy = 10^(-5);
elseif (nargin == 3)
  sdpAccuracy = seesawAccuracy/10;
end

% channel ETD in Heisenberg picture
% D: inputDimension -> decoderOutputDimension
% T: decoderOutputDimension -> encoderInputDimension
% E: encoderInputDimension -> inputDimension
eInDim = getOutputdim(noiseChannel);
dOutDim = getInputdim(noiseChannel);

% random start channels
encoderChannel = CHSChannel(CRandomChannel(eInDim, inDim));

% constraints
%%%%%%%%%%%%%%

% x = decoder in Heisenberg picture
%
% define A
% < iA | xChannel(| iI >< iJ |) | iB >
```

```
% = < iA | HSChannel(xChannel)(| iB >< iJ |) | iI >
% = HilbertSchmidtMap(xChannel)(iA,iI,iB,iJ)
%
% xChannel is unital: xChannel(1) = 1
% xChannel(1) = partialTrace_xInput(HilbertSchmidtMap(xChannel))
% = sum_{iI,iJ}( HilbertSchmidtMap(xChannel)(iA,iI,iB,iJ)*Δ(iI - iJ) )
Ad = zeros(dOutDim^2,inDim^2*dOutDim^2);
for iA =  1:dOutDim
  for iB = 1:dOutDim
    vtmp = zeros(dOutDim,inDim,dOutDim,inDim);
    vtmp(iA,:,iB,:) = eye(inDim); % Δ(iI-iJ)
    % reshape works columnwise:
    % index(array(a,i,b,j)) = a
    %                       + (i-1)*aMax
    %                       + (b-1)*aMax*iMax
    %                       + (j-1)*aMax*iMax*bMax
    Ad(iA+(iB-1)*dOutDim,:) = reshape(vtmp,[1,inDim^2*dOutDim^2]);
  end % iB
end % iA
clear vtmp
Ad = sparse(Ad);
% define b
% xChannel(1) = A*x = 1 = b
bd = reshape(eye(dOutDim),[dOutDim^2,1]);
% define cone
Kd.s = [inDim*dOutDim]; % mat(xChannel) ≥ 0
Kd.scomplex = 1; % mat(xChannel) is hermitian
Kd.ycomplex = [1:dOutDim^2]; % number of constraints (A*x)_i = b_i
% from xChannel(1) = 1

% x = encoder in Schroedinger picture
% define A
% < iA | xChannel(| iI >< iJ |) | iB >
% = < iA | HSChannel(xChannel)(| iB >< iJ |) | iI >
% = HilbertSchmidtMap(xChannel)(iA,iI,iB,iJ)
%
% xChannel is trace preserving: trace(xChannel(| iI >< iJ |)) = Δ(iI-iJ)
%
% trace(xChannel(| iI >< iJ |))
% = sum_{a,b}(HilbertSchmidtMap(xChannel)(iA,iI,iB,iJ)*Δ(iA-iB))
Ae = zeros(inDim^2,inDim^2*eInDim^2);
for iI = 1:inDim
  for iJ = 1:inDim
    vtmp = zeros(eInDim,inDim,eInDim,inDim);
    vtmp(:,iI,:,iJ) = eye(eInDim);
    % reshape works columnwise:
    % index(array(a,i,b,j)) = a
    %                       + (i-1)*aMax
    %                       + (b-1)*aMax*iMax
    %                       + (j-1)*aMax*iMax*bMax
    Ae(iI+(iJ-1)*inDim,:) = reshape(vtmp,[1,inDim^2*eInDim^2]);
```

```
  end %iJ
end % iI
clear vtmp
Ae = sparse(Ae);
% define b
% xChannel is trace preserving: trace(xChannel(| iI >< iJ |))
% = Δ(iI-iJ) = b
be = reshape(eye(inDim),[inDim^2,1]);
% define cone
Ke.s = [inDim*eInDim]; % mat(xChannel) ≥ 0
Ke.scomplex = 1; % mat(xChannel) is hermitian
Ke.ycomplex = [1:inDim^2]; % number of constraints (A*x)_i = b_i
% from trace preserving condition

% parameters
%%%%%%%%%%%%
pars.fid = 0; % quiet
pars.eps = sdpAccuracy;

% seesaw
fidelity = -1;
endSeesaw = false;
while not(endSeesaw)
  oldFidelity = fidelity;

  % decoder iteration
  etChannel = concatenate(encoderChannel, noiseChannel);
  f = reshape(getHilbertSchmidtMap(dual(etChannel)), ...
        [inDim*dOutDim, inDim*dOutDim]);
  c = - vec((f'))/inDim^2;
  % solve SDP
  [x,y,info] = sedumi(Ad,bd,c,Kd,pars);

  % decoder in Heisenberg picture
  decoderChannel = CHSChannel(reshape(x,[dOutDim,inDim,dOutDim,inDim]));

  % encoder iteration
  tdChannel = concatenate(noiseChannel, decoderChannel);
  f = reshape(getHilbertSchmidtMap(tdChannel), ...
        [inDim*eInDim, inDim*eInDim]);
  c = - vec((f'))/inDim^2;
  % solve SDP
  [x,y,info] = sedumi(Ae,be,c,Ke,pars);

  % compute fidelity
  fidelity = real(-(c'*x));
  % encoder in Heisenberg picture
  encoderChannel = dual(CHSChannel(reshape(x, ...
                      [eInDim,inDim,eInDim,inDim])));

  if abs(oldFidelity - fidelity) < seesawAccuracy;
```

```
    endSeesaw = true;
  end
end % while
```

The matrices `Ad` and `Ae` for the affine equality constraints that model the partial trace conditions are both sparse. The positive semidefinite cones are defined via `Kd` and `Ke` for the decoder and encoder, respectively. The channel fidelity is given via $f(\widetilde{T}) = \mathrm{tr}(\widetilde{F}\widetilde{T})$ with the choices $F = (ET)_*$ form equation (3.28) for the iteration of the decoder D, and $F = TD$ from equation (3.30) for the iteration of the encoder E_*. The method `getHilbertSchmidtMap` returns the array $t(a, i, b, j) = \langle a|\widetilde{T}(|b\rangle\langle j|)|i\rangle$, as defined in equation (3.84). The method `dual` maps T to T_* and vice versa. As one can see, most of the required data format transformations can be done via `reshape`, so they produce little computational overhead.

Alternatively, one could implement the parametrization of Audenaert and De Moor with implicit partial trace constraint and realize the positivity of the Jamiolkowsky dual in the dual conic program formulation (3.89). However, this would lead to an even larger matrix A, in particular, since SeDuMi uses the matrix cone

$$S = \left\{ x \in \mathbb{C}^{p^2} \,|x + x^* \geq 0 \right\}$$

for the dual cone $\mathcal{C}^*$. For example, in the Schrödinger picture, the implicit parametrization has $(d_1^2 - 1)(d_2^2 - 1) + (d_2^2 - 1)$ parameters. The matrix A^* would map these parameters to the corresponding Jamiolkowsky matrix, which has to be positive semidefinite. In order to have $x \geq 0$ in terms of the above matrix cone S, we would require that $x \in S$, $ix \in S$, and $-ix \in S$, so $x + x^* \geq 0$ and $x = x^*$. In total, A^* would be a $(3(d_1 d_2)^2) \times ((d_1^2 - 1)(d_2^2 - 1) + (d_2^2 - 1))$-matrix. On the other hand, if we were only interested in the optimal fidelity and not in the channel for that optimal fidelity, we could implement the simplified dual problem.

The main advantage of semidefinite programming is, that in the case of zero duality gap, we obtain a certified optimum up to any given accuracy. However, this comes at the price of computational overhead compared to the power iteration. Furthermore, we cannot restrict the encoding to be isometric, as restrictions on the rank of $\widehat{T}$ are not linear. On the other hand, we can combine both iteration methods. We take the iteration result of the power iteration, solve the dual problem, and show that the duality gap is zero to certify that we indeed found the global optimum up to the given accuracy. This way, we only need to solve the dual problem via semidefinite programming, which is computationally much cheaper than the primal problem.

3.3 Applications

The above iteration methods have the following characteristic features:

- For known channels these methods yield excellent results without requiring any special properties of the channels like symmetry.

- The optimization of either encoding or decoding is a semidefinite problem for which the solution is a certified global optimum. The process of alternatingly optimizing these therefore improves the objective in every step, and hence converges to a local optimum. However, there is no guarantee for having found the global optimum.

- The methods suffer from the familiar explosion of difficulty as the system size is increased. Correction schemes like the five qubit code can still be handled on a PC, but a nine qubit code would involve optimization over $2^{10} \times 2^{10}$-matrices, which is set up by multiplying and contracting some matrices in 2^{18} dimensions. This may be possible on large machines, but it is clear that these methods are useless for asymptotic questions.

- The power iteration has an advantage here, because it has a lower flop count and it works with a fixed number of Kraus operators. So one can restrict the encoding to be isometric, that is, to have a single Kraus operator, which turns out to be optimal in many cases. This cuts down on the problem size, at least for the optimization of encodings.

- For asymptotic coding theory one still needs codes which can be described also for very large dimensions, be it by explicit parametrization or by a characterization of typical instances of random codes. It is here that methods transferred from classical coding theory will continue to be useful.

3.3.1 Test Cases

We can use noise situations, where the optimal solution is already known, to create test cases for the seesaw iteration as well as for the optimizations of the linear objectives in the encoder and decoder iteration steps.

3.3.1.1 Noiseless Subsystems

As the first test case we use a noisy channel, where the algebra generated by the error operators has a noiseless subsystem [71]. We expect that the seesaw iteration finds a code (E, D) that encodes into this subsystem, and thus allows perfect transmission of quantum states, leading to a channel fidelity equal to one.

Consider the channel $T\colon \mathcal{B}(\mathcal{H}) \to \mathcal{B}(\mathcal{H})$, $T(\rho) = \sum_\alpha t_\alpha \rho\, t_\alpha^*$. Following Zanardi [72], we define the interaction algebra $\mathcal{A}$ as the algebra generated by the Kraus operators t_α. The algebra $\mathcal{A}$ can be written as a direct sum of complex square matrix algebras

$\mathcal{M}_{d_J}$ with multiplicity n_J,

$$\mathcal{A} \cong \bigoplus_{J\in\mathcal{J}} \mathbb{1}_{n_J} \otimes \mathcal{M}_{d_J}, \tag{3.93}$$

where $\mathcal{J}$ labels the irreducible components of $\mathcal{A}$. The factors $\mathbb{1}_{n_J}$ are then called noiseless subsystems. The state space decomposition that corresponds to (3.93) is[14]

$$\mathcal{H} \cong \bigoplus_{J\in\mathcal{J}} \mathbb{C}^{n_J} \otimes \mathbb{C}^{d_J}. \tag{3.94}$$

If there even exists an irreducible representation J of $\mathcal{A}$ with $d_J = 1$ and $n_J \geq 2$, we can just encode a quantum state $|\psi\rangle$ into the corresponding summand in (3.94). The encoded quantum state is then, up to a global phase, preserved under the action of the noise. Such a summand is called *decoherence free subspace*.

We will now see how to read off the available quantum codes from the decomposition (3.93). To this end, we look at the commutant $\mathcal{A}'$ of $\mathcal{A}$ defined by,

$$\mathcal{A}' = \{X \in \mathcal{B}(\mathcal{H}) \,|\, XA = AX \text{ for all } A \in \mathcal{A}\}.$$

From equation (3.93) we get

$$\mathcal{A}' \cong \bigoplus_{J\in\mathcal{J}} \mathcal{M}_{n_J} \otimes \mathbb{1}_{d_J}. \tag{3.95}$$

Thus we can project to each summand of the direct sum by projectors $Q_J = \mathbb{1}_{n_J} \otimes \mathbb{1}_{d_J} \in \mathcal{A} \cap \mathcal{A}'$. Furthermore, we can use the projected space $Q_J\mathcal{H}$ as range for an isometric encoding with isometry v. Let $\{|J,\mu,\nu\rangle\}$ be an orthonormal basis associated with the decomposition (3.94), i. e., $\mu = 1,\ldots,n_J$ and $\nu = 1,\ldots d_j$. Then, we fix J and ν and set

$$v = \sum_{\mu} |J\mu\nu\rangle\langle\mu|. \tag{3.96}$$

From (3.93) it follows that

$$\begin{aligned} v^* t_\alpha^* t_\beta v &= \sum_{\mu=1}^{n_J}\sum_{\xi=1}^{n_J} |\mu\rangle\langle J\mu\nu|t_\alpha^* t_\beta|J\xi\nu\rangle\langle\xi| \\ &= \sum_{\mu=1}^{n_J}\sum_{\xi=1}^{n_J} |\mu\rangle\langle J\mu\nu|\mathbb{1}\otimes X_{\alpha\beta}|J\xi\nu\rangle\langle\xi| \\ &= \sum_{\mu=1}^{n_J}\sum_{\xi=1}^{n_J} \delta_{\mu,\xi}\ \lambda^{J,\nu}_{t_\alpha^* t_\beta}|\mu\rangle\langle\xi| = \lambda^{J,\nu}_{t_\alpha^* t_\beta}\mathbb{1}, \end{aligned} \tag{3.97}$$

where $X_{\alpha\beta}$ is some operator and $\lambda^{J,\nu}_{t_\alpha^* t_\beta}$ is a complex number, both depending on the noise operators. Hence, the isometry v satisfies the Knill-Laflamme condition (3.2)

[14] Note that each summand of the direct sum is a Hilbert space split into good and bad as in Keyl's error corrector's dream [51].

on page 26, and therefore it is an encoding isometry of a perfect error correcting code. The code dimension is n_J, so we require $n_J \geq 2$, which amounts to having a noncommutative commutant $\mathcal{A}'$ in (3.95).[15]

Observe that there is a severe limitation of this algebraic framework. The interaction algebra $\mathcal{A}$ for a channel T contains all possible products of Kraus operators, and thus equation (3.97) holds for any such a product. On the other hand, the Knill-Laflamme condition (3.2) of Theorem 3.1.4 does not at all require to hold for any product, but only for any pair of Kraus operators. For example, we look at a system that is the fivefold tensor product of qubit systems, $\mathcal{H} = (\mathbb{C}^2)^{\otimes 5}$. We take the error operators e_α to be the tensor products

$$e_\alpha = e_{\alpha,1} \otimes ... \otimes e_{\alpha,5}, \tag{3.98}$$

where at most one factor $e_{\alpha,i}$, $i = 1, \ldots, 5$, is a Pauli operator and all others are the identity. In total, we have 16 error operators, one is the identity $\mathbb{1}^{\otimes 5}$, and additional $3 \cdot 5$ operators as we have all combinations of three Pauli operators $\sigma_x, \sigma_y, \sigma_z$ on five tensor positions. In order that there exists a perfect error correction code, the Knill-Laflamme condition (3.2) must hold for all combinations $e_\alpha^* e_\beta$, that is, for all operators

$$e_{\alpha,1}^* e_{\beta,1} \otimes ... \otimes e_{\alpha,5}^* e_{\beta,5}$$

with at most two factors $e_{\alpha,i}^* e_{\beta,i}$ different from the identity. However, the error operators (3.98) generate the full algebra $\mathcal{B}(\mathcal{H})$, and therefore the commutant of the interaction algebra is $\mathbb{C}\mathbb{1}$. This implies that there is no perfect quantum error correcting code of the form (3.96). Nevertheless, there exists a perfect quantum error correcting code for the single side errors (3.98), which is the famous five bit stabilizer code [57, 58].

Yet, there are cases in which a nontrivial noiseless subsystem exists. As test case, we choose the 3-qubit collective rotation channel,

$$T(A) = t_x A t_x + t_y A t_y + t_z A t_z,$$

where the Kraus operators are given by

$$t_\alpha = \frac{1}{\sqrt{3}} e^{i(\sigma_\alpha \otimes \mathbb{1} \otimes \mathbb{1} + \mathbb{1} \otimes \sigma_\alpha \otimes \mathbb{1} + \mathbb{1} \otimes \mathbb{1} \otimes \sigma_\alpha)},$$

with the Pauli operators σ_α. The t_α are self-adjoint, so T is a bistochastic channel, i. e., $T(\mathbb{1}) = T_*(\mathbb{1}) = \mathbb{1}$. Each Kraus operator is a collective Pauli rotation on all three qubits. For this channel, Holbrook *et al.* [73] show that the commutant of the interaction algebra is

$$\mathcal{A}' \cong (\mathbb{1}_2 \otimes \mathcal{M}_2) \oplus \mathbb{C}\mathbb{1}_4.$$

[15] For a more detailed explanation and further properties of noiseless subsystems, please refer to Zanardi [72].

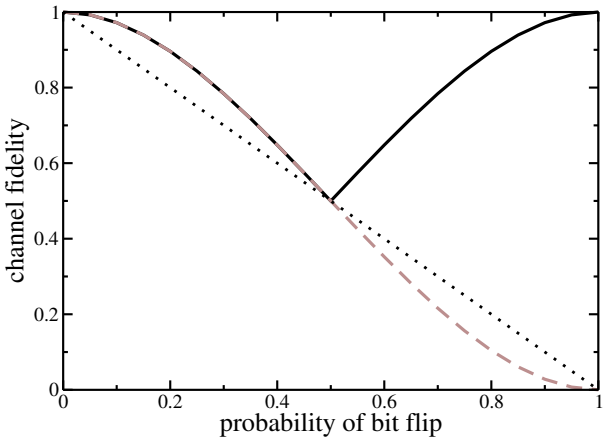

Figure 3.6: Comparison of the channel fidelity of no error correction (dotted line), the majority vote stabilizer (eq. (3.102), dashed line) and the iteration (solid line) applied to the 3-fold tensor product of the bit flip channel.

Thus, according to (3.97) there exists an isometry $v\colon \mathbb{C}^2 \to (\mathbb{C}^2)^{\otimes 3}$ that encodes a qubit into the noiseless subsystem $\mathcal{M}_2$. Using the decoder $D(A) = vAv^*$ we obtain $f_c(ETD) = f_c(\mathrm{id}_{\mathbb{C}^2}) = 1$ for the channel fidelity.

We use the collective rotation channel as test case for the seesaw iteration using either the subchannel power iteration or the semidefinite program for the optimizations of the encoder $E\colon \mathcal{B}((\mathbb{C}^2)^{\otimes 3}) \to \mathcal{B}(\mathcal{K})$ and the decoder $D\colon \mathcal{B}(\mathcal{K}) \to \mathcal{B}((\mathbb{C}^2)^{\otimes 3})$. The optimal channel fidelities found by the iteration are summarized in the following table:

	optimal channel fidelities found by the seesaw iteration with	
$\dim \mathcal{K}$	subchannel power iteration	semidefinite program
2	1.00	1.00
3	0.98	0.98
4	0.89	0.89
5	0.67	0.67
6	0.50	0.50
7	0.41	0.41
8	0.34	0.34

Both methods, power iteration and semidefinite program, yield the same channel fidelity in all cases. In particular, both find a perfect error correcting scheme in the qubit case, where a decoherence free subspace exists.

3.3.1.2 Bit-Flip Channel

Another test case for the seesaw iteration is given by the 3-fold tensor product of the bit flip channel as noise. The bit flip channel applies the Pauli matrix σ_x with probability p and does nothing with probability $(1-p)$,

$$T_p(\rho) = p\ \sigma_x \rho \sigma_x + (1-p)\rho. \tag{3.99}$$

We compare the result with the majority vote stabilizer code. This stabilizer code maps $|0\rangle \mapsto |000\rangle$ and $|1\rangle \mapsto |111\rangle$, i. e.,

$$E_3(\rho) = v\rho v^*, \tag{3.100}$$

with the isometry $v = |000\rangle\langle 0| + |111\rangle\langle 1|$. The usual Kraus decomposition of the noise $T_p^{\otimes 3}$ uses operators which are a combination of σ_x and $\mathbb{1}$ on the tensor factors, e.g., there is a Kraus operator $\sqrt{p^2(1-p)}(\sigma_x \otimes \sigma_x \otimes \mathbb{1})$. One says that this operator flips the first two qubits with probability $p^2(1-p)$. The majority vote decoder maps the subspaces, the Kraus operators with at most one tensor factor different from the identity maps to, back to the undisturbed subspace, so they are corrected,

$$D_3(\rho) = \sum_{\alpha=1}^{4} d_\alpha \rho d_\alpha^*, \tag{3.101}$$

$$\begin{aligned} d_1 &= |0\rangle\langle 000| + |1\rangle\langle 111|, \\ d_2 &= |0\rangle\langle 001| + |1\rangle\langle 110|, \\ d_3 &= |0\rangle\langle 010| + |1\rangle\langle 101|, \\ d_4 &= |0\rangle\langle 100| + |1\rangle\langle 011|. \end{aligned}$$

Of course this is good for rare bit flip occurrences and the code makes things worse at bit flip probabilities greater than $1/2$. The bit flip is the only quantum error with a classical analogue. Therefore we expect the iteration to find the known classical majority vote code, with an additional flip for $p > 1/2$.

The channel fidelity using the majority vote stabilizer is

$$f_c(D_3 T_p^{\otimes 3} E_3) = (1-p)^3 + 3(1-p)^2 p. \tag{3.102}$$

This can be seen from the Kraus operators of the channel $D_3 T_p^{\otimes 3} E_3$ and equation (3.5) for the channel fidelity, $f_c(T) = 1/d^2 \sum_\alpha |\operatorname{tr}(t_\alpha)|^2$. The Kraus operator of $T_p^{\otimes 3}$ that is proportional to the identity is perfectly corrected by the code (E_3, D_3), as well as all Kraus operators proportional to a bit flip on a single tensor factor only. The identity has probability $(1-p)^3$, a single side error has the total probability of $3(1-p)^2 p$. For all other Kraus operators of $T_p^{\otimes 3}$, the majority vote results in the wrong results, and therefore these errors correspond to Kraus operators of $D_3 T_p^{\otimes 3} E_3$

proportional to a bit flip. Since $\operatorname{tr}\sigma_x = 0$, the terms of the channel fidelity are zero in this case. In summary we get the fidelity (3.102).

The iteration is started with random channels as encoder and decoder and it stops if the difference between the fidelities of sequent iteration steps is below a given threshold. The results are shown in Figure 3.6. The fidelity of the best code (E, D) found by the seesaw iteration coincides with the majority rule code within numerical accuracy. Like the stabilizer, the iterated encoder is isometric, that is, it has only one Kraus operator. Also the decoder in both cases is homomorphic, i.e. it can be written in the form $D(x) = u(x \otimes \mathbb{1})u^*$ in the Heisenberg picture, where u is unitary. Moreover, one can find a unitary transformation on $(\mathbb{C}^2)^{\otimes 3}$ mapping (E, D) to the majority rule code. Near the crossover point only the upper branch (solid line) produces a stable solution in the (sub)channel power iteration: if one perturbs the exact expression for the lower branch (dashed line) by mixing it with a small fraction of a random channel, the iteration again finds the upper branch. Thus the 3-fold tensor product of the bit flip channel also provides a test case for the treatment of local but not global optimal codes.

Together, these test cases cover the situations noiseless subspace, stabilizer code, and local optimal solution that is not global optimal. All are suitable for test automation as well as unit tests, given either encoder or decoder is fixed to the optimal solution while iterating the other.

3.3.2 Depolarizing Channel

We will now take the five-fold tensor product of the depolarizing qubit channel as noise. We will compare the optimal code obtained from the seesaw iteration with the five bit code [57, 58]. The performance of the five bit code in this case was our example on page 28. The depolarizing channel, given by equation (3.6),

$$T_p(A) = p \operatorname{tr}\left(A \frac{1}{d}\mathbb{1}\right)\mathbb{1} + (1-p)\,A,$$

can be interpreted as a channel that replaces the input state by the complete mixed state $\rho = \mathbb{1}/d$ with probability p and leaves the input system untouched with probability $(1-p)$. Choosing the depolarizing channel as noise complies with a worst case analysis, as every channel can be turned into a depolarizing channel with the same channel fidelity via a twirl operation [74].

From the Kraus decomposition (3.7) we see that T_p is a channel even for $1 \leq p \leq 4/3$. The three Pauli operators are applied with equal probability, that is, no error dominates. This can be interpreted by looking at the effect of T_p on the Bloch sphere. The completely mixed state corresponds to the center of the Bloch sphere, so for $p \leq 1$ the effect of the depolarizing channel on the Bloch sphere is a uniform contraction. For $1 \leq p \leq 4/3$, we get a contracted inversion of the sphere, the

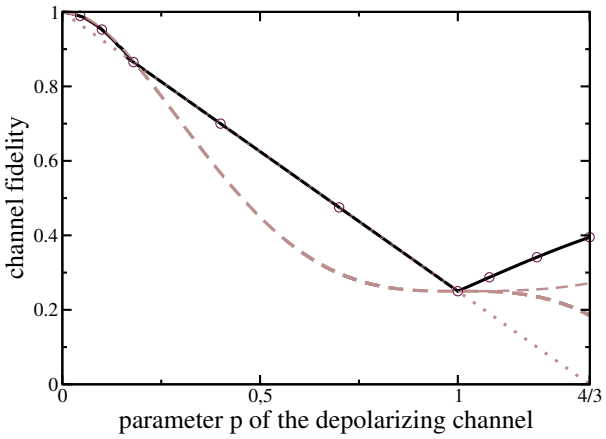

Figure 3.7: Comparison of the channel fidelity of no error correction (dotted line), five bit code (dashed line) and the iteration (solid line) applied to the 5-fold tensor product of depolarizing channel with parameter p. Note that T is a channel even for $p > 1$. For $p > 1$, the upper dashed line shows the channel fidelity of the five bit encoder and optimal decoder, while the lower dashed line corresponds to the usual five bit code. The circles correspond to results of the seesaw iteration with a restriction to isometric encodings.

channel T_p approximates the Universal-NOT gate $|\varphi\rangle \mapsto |\varphi^\perp\rangle \forall \varphi$[16]. Note, however, that a perfect inversion of a Bloch sphere, or equivalently a perfect Universal-NOT gate, is not a physical operation.

We will compare the iteration result to the five bit code (E_5, D_5), where we already know from equation (3.8) on page 28 that it is better than no coding for small values of p. No coding means that a single channel T_p is used, or equivalently, that the encoder maps the input system to one of the tensor factors and the decoder simply traces out the remaining four qubits. The five bit code is the smallest stabilizer that corrects localized errors. Therefore we expect to find better codings for larger values of p, where the local errors are no longer the dominant Kraus operators.

We use the seesaw iteration with the subchannel power iteration for the single optimizations. The iteration results reported below were computed by starting from various random initial configurations. The iteration was stopped when the gain of fidelity was below some threshold. The results are shown in Figure 3.7. For the parameter range $0 \leq p \leq 1 - \sqrt{2/3} \approx 0.18$, the iteration does not find any code that is better than the five bit code. For $1 - \sqrt{2/3} \leq p \leq 1$, we find no code that is better than no coding at all. That we do not find any better codes even near the crossover point between the five bit code and no coding is very surprising in view of the fact that the five bit code is not at all designed to give good results for large

[16] See [75] for approximations to the Universal-NOT gate.

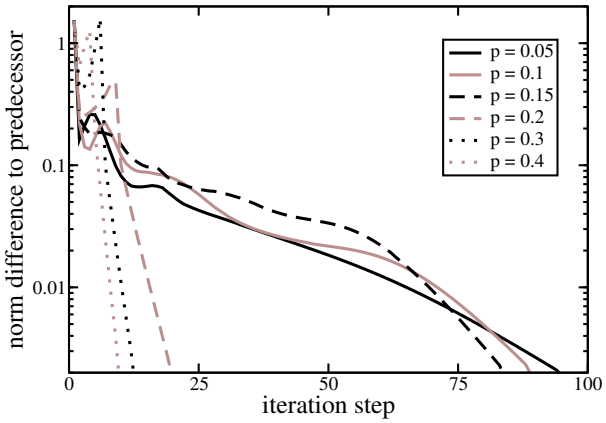

Figure 3.8: Norm difference between successive encoders in the iteration for the 5-fold tensor product of the depolarizing channel with depolarization probability p.

errors. However, for $1 \leq p \leq 4/3$, the iteration result is clearly superior to the five bit code, even if we just take the five bit encoding and optimize the decoder with the subchannel power iteration (upper dashed line in Figure 3.7).

To analyze the convergence of the subchannel power iteration, we look at the operator norm differences of the successive encoders and decoders. The norm difference is taken as

$$\|\widehat{E}_n - \widehat{E}_{n-1}\|_2,$$

where $\|\cdot\|_2$ denotes the operator norm and $\widehat{E}$ is the matrix associated with the channel E, or equivalently $\widetilde{E}$, according to equation (3.15) on page 31. The norm differences for various parameters p are shown in Figure 3.8, the plot for the decoders looks similar. We see that the final phase of the iteration is typically an exponential approach to the fixed point. Also, the convergence in the cases where no coding is optimal is much faster than the convergence in the cases where the five bit code stabilizer is optimal. The iteration results turn out to have full support, that is, the result was always a channel, not a subchannel. Furthermore, we did not find any stable suboptimal fixed point.

Figure 3.7 also shows the results of the seesaw iteration where we used the semidefinite program for the decoder iteration and the subchannel power method restricted to a single Kraus operator for the encoder iteration (circles in the figure). That is, we restrict E to be an isometric encoding. This shows that isometric encoding is also optimal for the case $p > 1$.

We now have a closer look at the case $p > 1$. From the Kraus decomposition of the depolarizing channel (3.7), we see that we can improve the channel fidelity for the depolarizing channel T_p in this case using a unitary channel with one of the Pauli matrices as decoder, e.g., with the code $(E = \mathrm{id}, D(A) = XAX)$. From equation

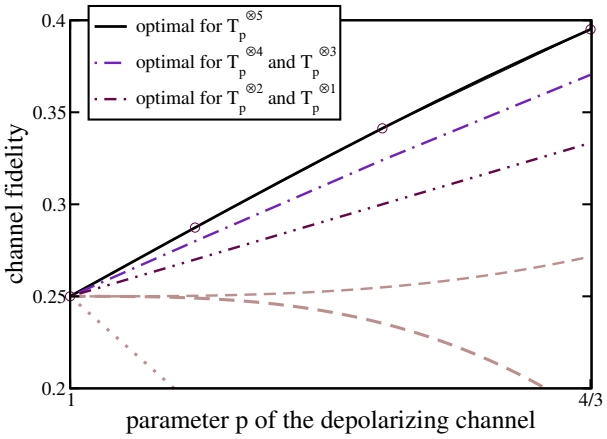

Figure 3.9: Comparison of the optimal codes for different number of qubits found by the seesaw iteration for the depolarizing channel with $1 \leq p \leq 4/3$. It shows that the code found in the five qubit case is better than the code found for four or three qubits. The optimal channel fidelity found for $T_p^{\otimes 2}$ and T_p is $p/4$, which is equal to the channel fidelity of a code that simply rotates back one of the Pauli operators, e.g., $E(A) = A$, $D(A) = XAX$.

(3.5), $f_c(T) = 1/d^2 \sum_\alpha |\operatorname{tr}(t_\alpha)|^2$, we immediately obtain that the channel fidelity for such a code is $p/4$. Figure 3.9 shows the seesaw iteration results for $T_p^{\otimes n}$ as noise for the cases $n = 1, \ldots 5$. As shown in the figure, there exists a code with fidelity larger than $p/4$ already in the three qubit case $n = 3$. The plot also indicates that the code obtained in the five qubit case $n = 5$ uses all five qubits and, unlike the case $n = 4$, cannot be reduced to a code on a smaller system.

In summary, we have found a new code that uses all five qubits and is superior to the five bit stabilizer code for $p > 1$. Interestingly, we can restrict the encoding operation for the new code to be isometric, which is a basic feature of Knill-Laflamme type encodings. For $p < 1$ we haven't found any code that outperforms existing codes. Up to about 18 percent depolarization probability, the five-bit code is optimal. For larger depolarization probabilities, the best way of using the fivefold tensor product of the depolarizing channel is to do no coding, which means to just use one of the tensor factors T_p. Furthermore, for the smaller system $T_p^{\otimes 4}$ and $p < 1$, the iteration indicates that no coding is optimal even for small p. However, this has little bearing on general channels, since the depolarizing channel is highly symmetric. If we modify the depolarizing channel such that the input state is, with probability p, replaced by $|1\rangle\langle 1|$ instead of the completely mixed state $1/2\mathbb{1}$, the seesaw iteration always finds a code with better performance than the stabilizer code for the complete range $0 < p < 1$. This suggests to investigate less symmetric noise as done in the following.

3.3.3 Amplitude Damping Channel

We now take the four-fold tensor product of the amplitude damping channel as noise and consider the encoding of qubit systems. This case is interesting, because Leung *et al.* [14] developed an approximate error correcting code for this case. Their code does not satisfy the Knill-Laflamme condition of Theorem 3.1.4. Note that approximate quantum error correction has no classical analogue. Also, their code is not based on classical coding techniques. In particular, they do not use a decomposition into Pauli matrices. Furthermore, their code violates the quantum singleton bound (12.4.3 [35]), which states that the smallest code that corrects single qubit errors must at least use five qubits to encode one qubit. In summary this means that the four-fold tensor product of the amplitude damping channel corresponds to a purely quantum setting.

The amplitude damping channel for qubits is defined as $T_\gamma \colon \mathcal{B}(\mathbb{C}^2) \to \mathcal{B}(\mathbb{C}^2)$,

$$T_\lambda(A) = t_{0,\lambda}^* A\, t_{0,\lambda} + t_{1,\lambda}^* A\, t_{1,\lambda}, \tag{3.103}$$

where the Kraus operators are given by

$$\begin{aligned} t_{0,\lambda} &= \begin{pmatrix} 1 & 0 \\ 0 & \sqrt{1-\gamma} \end{pmatrix} = \frac{1}{2}\Big(\big(1+\sqrt{1-\gamma}\big)\mathbb{1} + \big(1-\sqrt{1-\gamma}\big)\sigma_z \Big), \\ t_{1,\lambda} &= \begin{pmatrix} 0 & \sqrt{\gamma} \\ 0 & 0 \end{pmatrix} = \frac{\sqrt{\gamma}}{2}\Big(\sigma_x + i\sigma_y\Big). \end{aligned} \tag{3.104}$$

Here σ_x, σ_y, and σ_z are the corresponding Pauli operators and the damping parameter γ is the probability to flip the state $|1\rangle$ to the state $|0\rangle$. These channels form a semigroup $T_\gamma T_\eta = T_{\gamma\eta}$. They can be used to model energy dissipation. For example, γ can be thought of as the probability of losing a photon (8.3.5 [35]). From equation (3.104) we see that a Pauli decomposition would involve all four matrices of the Pauli basis, so a stabilizer code for the correction of local errors, i. e., small values of γ, would indeed need to encode a logical qubit into five qubits. The channel fidelity of the amplitude damping channel is

$$f_c(T_\gamma) = \frac{1}{d^2}\sum_\alpha |\operatorname{tr}(t_\alpha)|^2 = \frac{1}{4}\left|\operatorname{tr}\left(\frac{1}{2}\left(1+\sqrt{1-\gamma}\right)\mathbb{1}\right)\right|^2 = \left(\frac{1+\sqrt{1-\gamma}}{2}\right)^2, \tag{3.105}$$

where we used the fact that the Pauli matrices are traceless in the second equality.

We will compare the results of the iteration with the four bit code by Leung *et al.* [14]. The four bit code has the isometric encoder

$$E_4(A) = v^* A v, \tag{3.106}$$

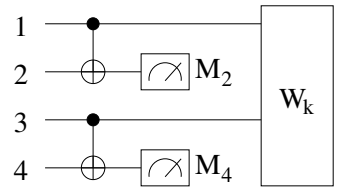

Figure 3.10: Syndrome measurement of the four bit amplitude damping code [14].

where $v = |0_L\rangle\langle 0| + |1_L\rangle\langle 1|$, with the logical states

$$\begin{aligned} |0_L\rangle &= \frac{1}{\sqrt{2}}\left(|0000\rangle + |1111\rangle\right) \\ |1_L\rangle &= \frac{1}{\sqrt{2}}\left(|0011\rangle + |1100\rangle\right). \end{aligned} \tag{3.107}$$

To see that this code does not satisfy the Knill-Laflamme condition (3.2), we consider the error operator $t_{0,\lambda}^{\otimes 4}$ of the four-fold tensor product of the amplitude damping channel. The Knill-Laflamme condition requires that

$$\langle 0_L|(t_{0,\lambda}^{\otimes 4})^* t_{0,\lambda}^{\otimes 4}|0_L\rangle = \langle 1_L|(t_{0,\lambda}^{\otimes 4})^* t_{0,\lambda}^{\otimes 4}|1_L\rangle.$$

In contrast, we have

$$\begin{aligned} \langle 0_L|(t_{0,\lambda}^{\otimes 4})^* t_{0,\lambda}^{\otimes 4}|0_L\rangle &= \frac{1}{2}\left(1 + (1-\gamma)^4\right) = 1 - 2\gamma + 3\gamma^2 - 2\gamma^3 + \frac{1}{2}\gamma^4 \\ \langle 1_L|(t_{0,\lambda}^{\otimes 4})^* t_{0,\lambda}^{\otimes 4}|1_L\rangle &= (1-\gamma)^2 = 1 - 2\gamma + \gamma^2. \end{aligned}$$

Thus, the Knill-Laflamme does not hold exactly, but only to the first order in γ.

The decoder was constructed by distinguishing the possible outcomes of a pure input qubit and applying the appropriate correction procedure. The distinguishing of possible outcomes, also known as syndrome measurement, is done as shown in Figure 3.10. The four lines correspond to the four qubits,  to the C-NOT gate[17], and is a σ_z-measurement, i. e., the projection onto $|0\rangle$ or $|1\rangle$. The circuit W_k corresponds to the correction procedure, chosen according to the measurement results from the two meters $k = (M_2, M_4)$. They are shown in Figure 3.11. Here  depicts the NOT gate, which is equivalent to a σ_x unitary operation. The parameters θ and θ' of the rotation and controlled-rotation gates depend on the amplitude damping parameter γ. They are give by the equations $\tan\theta = (1-\gamma)^2$ and $\cos\theta' = 1 - \gamma$. The rotation itself is defined as $U_\Theta = \exp(i\Theta\sigma_y)$, where σ_y is the corresponding Pauli matrix. The controlled-rotation only applies the rotation to the target qubit $|t\rangle$ if the control qubit $|c\rangle$ is set[18],

$$|c\rangle \otimes |t\rangle \mapsto |c\rangle \otimes U^c|t\rangle.$$

[17] See the textbook [35] for a complete overview of the quantum circuit model.
[18] See chapter 4.3 in [35] for a detailed discussion of controlled operations.

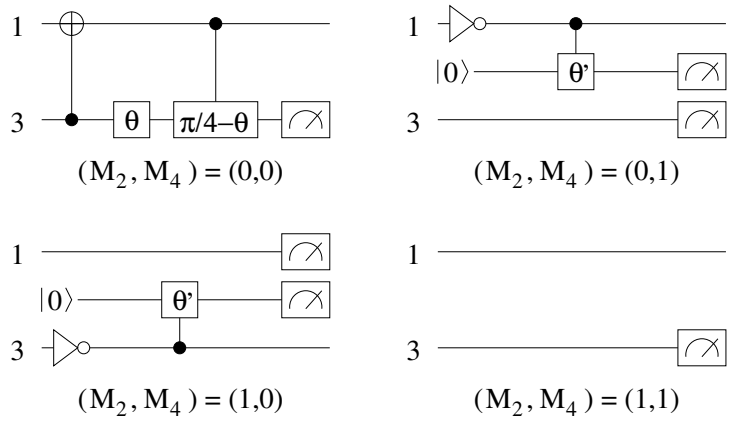

Figure 3.11: Correction procedures of the four bit amplitude damping code [14]. These are the circuits W_k in Figure 3.10 for all possible tuples of $k = (M_2, M_4)$. The angles θ and θ' of the rotation and controlled-rotation gates are specified by the amplitude damping parameter γ through $\tan\theta = (1-\gamma)^2$ and $\cos\theta' = 1-\gamma$. Given the angle Θ, the unitary of the rotation and controlled-rotation is $\exp(i\Theta\sigma_y)$.

The control qubit is marked with a black dot in the Figure. The gate $W_{1,1}$ is not specified in [14], as that syndrome does not occur with a probability $\mathcal{O}(\gamma)$ or above. Here we take it to be a σ_z-measurement of the third qubit.

Note that the circuits $W_{0,1}$ and $W_{1,0}$ use an additional ancilla qubit prepared in the $|0\rangle$ state, so both circuits translate to four Kraus operators. Together with the other two circuits we have 12 Kraus operators in total. Apart from $W_{1,1}$, all Kraus operators are parametrized with the amplitude damping parameter γ. The Kraus operators of the decoder

$$D_{4,\gamma}(A) = \sum_{\alpha=1}^{12} d^*_{\alpha,\gamma} A d_{\alpha,\gamma} \tag{3.108}$$

are

$$d_1^* = \begin{pmatrix} 0 & 0 \\ 0 & 0 \\ 0 & 0 \\ 0 & 0 \\ 0 & 0 \\ 1 & 0 \\ 0 & 0 \\ 0 & 0 \\ 0 & 0 \\ 0 & 1 \\ 0 & 0 \\ 0 & 0 \\ 0 & 0 \\ 0 & 0 \\ 0 & 0 \\ 0 & 0 \end{pmatrix}, \qquad d_2^* = \begin{pmatrix} 0 & 0 \\ 0 & 0 \\ 0 & 0 \\ 0 & 0 \\ 0 & 0 \\ 0 & 0 \\ 1 & 0 \\ 0 & 0 \\ 0 & 0 \\ 0 & 0 \\ 0 & 1 \\ 0 & 0 \\ 0 & 0 \\ 0 & 0 \\ 0 & 0 \\ 0 & 0 \end{pmatrix}, \qquad d_{3,\gamma}^* = \begin{pmatrix} \frac{1}{\sqrt{1+(\gamma-1)^4}} & 0 \\ 0 & 0 \\ 0 & 0 \\ 0 & \frac{1}{\sqrt{2}} \\ 0 & 0 \\ 0 & 0 \\ 0 & 0 \\ 0 & 0 \\ ou0 & 0 \\ 0 & 0 \\ 0 & 0 \\ 0 & 0 \\ 0 & \frac{1}{\sqrt{2}} \\ 0 & 0 \\ 0 & 0 \\ \frac{(\gamma-1)^2}{\sqrt{1+(\gamma-1)^4}} & 0 \end{pmatrix},$$

$$d_{4,\gamma}^* = \begin{pmatrix} -\frac{(\gamma-1)^2}{\sqrt{1+(\gamma-1)^4}} & 0 \\ 0 & 0 \\ 0 & 0 \\ 0 & \frac{1}{\sqrt{2}} \\ 0 & 0 \\ 0 & 0 \\ 0 & 0 \\ 0 & 0 \\ 0 & 0 \\ 0 & 0 \\ 0 & 0 \\ 0 & 0 \\ 0 & -\frac{1}{\sqrt{2}} \\ 0 & 0 \\ 0 & 0 \\ \frac{1}{\sqrt{1+(\gamma-1)^4}} & 0 \end{pmatrix}, \quad d_{5,\gamma}^* = \begin{pmatrix} 0 & 0 \\ 0 & 1-\gamma \\ 0 & 0 \\ 0 & 0 \\ 0 & 0 \\ 0 & 0 \\ 0 & 0 \\ 0 & 0 \\ 0 & 0 \\ 0 & 0 \\ 0 & 0 \\ 0 & 0 \\ 0 & 0 \\ 0 & 0 \\ 1 & 0 \\ 0 & 0 \\ 0 & 0 \end{pmatrix}, \quad d_{6,\gamma}^* = \begin{pmatrix} 0 & 0 \\ 0 & -\sqrt{(2-\gamma)\gamma} \\ 0 & 0 \\ 0 & 0 \\ 0 & 0 \\ 0 & 0 \\ 0 & 0 \\ 0 & 0 \\ 0 & 0 \\ 0 & 0 \\ 0 & 0 \\ 0 & 0 \\ 0 & 0 \\ 0 & 0 \\ 0 & 0 \\ 0 & 0 \\ 0 & 0 \end{pmatrix},$$

$$d^*_{7,\gamma} = \begin{pmatrix} 0 & 0 \\ 0 & 0 \\ 0 & 1-\gamma \\ 0 & 0 \\ 0 & 0 \\ 0 & 0 \\ 0 & 0 \\ 0 & 0 \\ 0 & 0 \\ 0 & 0 \\ 0 & 0 \\ 0 & 0 \\ 0 & 0 \\ 0 & 0 \\ 1 & 0 \\ 0 & 0 \end{pmatrix}, \qquad d^*_{8,\gamma} = \begin{pmatrix} 0 & 0 \\ 0 & 0 \\ 0 & -\sqrt{(2-\gamma)\gamma} \\ 0 & 0 \\ 0 & 0 \\ 0 & 0 \\ 0 & 0 \\ 0 & 0 \\ 0 & 0 \\ 0 & 0 \\ 0 & 0 \\ 0 & 0 \\ 0 & 0 \\ 0 & 0 \\ 0 & 0 \end{pmatrix}, \quad d^*_{9,\gamma} = \begin{pmatrix} 0 & 0 \\ 0 & 0 \\ 0 & 0 \\ 0 & 0 \\ 0 & 1-\gamma \\ 0 & 0 \\ 0 & 0 \\ 1 & 0 \\ 0 & 0 \\ 0 & 0 \\ 0 & 0 \\ 0 & 0 \\ 0 & 0 \\ 0 & 0 \\ 0 & 0 \end{pmatrix},$$

$$d^*_{10,\gamma} = \begin{pmatrix} 0 & 0 \\ 0 & 0 \\ 0 & 0 \\ 0 & 0 \\ 0 & -\sqrt{(2-\gamma)\gamma} \\ 0 & 0 \\ 0 & 0 \\ 0 & 0 \\ 0 & 0 \\ 0 & 0 \\ 0 & 0 \\ 0 & 0 \\ 0 & 0 \\ 0 & 0 \\ 0 & 0 \\ 0 & 0 \end{pmatrix}, \quad d^*_{11,\gamma} = \begin{pmatrix} 0 & 0 \\ 0 & 0 \\ 0 & 0 \\ 0 & 0 \\ 0 & 0 \\ 0 & 0 \\ 0 & 0 \\ 0 & 0 \\ 0 & 1-\gamma \\ 0 & 0 \\ 0 & 0 \\ 1 & 0 \\ 0 & 0 \\ 0 & 0 \\ 0 & 0 \\ 0 & 0 \end{pmatrix}, \qquad d^*_{12,\gamma} = \begin{pmatrix} 0 & 0 \\ 0 & 0 \\ 0 & 0 \\ 0 & 0 \\ 0 & 0 \\ 0 & 0 \\ 0 & 0 \\ 0 & 0 \\ 0 & -\sqrt{(2-\gamma)\gamma} \\ 0 & 0 \\ 0 & 0 \\ 0 & 0 \\ 0 & 0 \\ 0 & 0 \\ 0 & 0 \\ 0 & 0 \end{pmatrix}.$$

The channel fidelity of the four bit code for the amplitude damping noise is[19]

$$f_c(E_4 T_\gamma^{\otimes 4} D_{4,\gamma}) = \frac{1}{2} + \frac{\sqrt{1+(\gamma-1)^4}}{2\sqrt{2}} + \gamma - \frac{\sqrt{1+(\gamma-1)^4}}{2\sqrt{2}}\gamma - \frac{15}{4}\gamma^2 + \frac{7}{2}\gamma^3 - \gamma^4.$$

Thus, the amplitude damping code is better than no coding for damping probabilities below $p \approx 0.25$.

We now compare the above code with numerical optimizations. The numerical results are obtained using the seesaw iteration with the semidefinite program approach for the single optimizations.

First, we look at the four qubit case, that is, the noise is $T_\gamma^{\otimes 4}$. The results are shown in Figure 3.12. In the figure, the channel fidelity of different quantum codes (E, D) are compared to the channel fidelity of no coding, that is, to that of the single use of the amplitude damping channel $f_c(T_\gamma)$ given by equation (3.105). The approximate error correcting four qubit code [14] does increase the channel fidelity for $0 < \gamma < 0.25$. For a larger value of the amplitude damping parameter, using the code is actually worse than no coding at all. If we fix the decoding to that of the four qubit code $D_{4,\gamma}$ and optimize the encoder ("encoder optimization" in the figure), we get an improved channel fidelity compared to that of the single amplitude damping channel for $0 < \gamma < 0.35$. Conversely, if we fix the encoder to E_4 and optimize the decoder channel ("decoder optimization" in the figure), the resulting code improves the channel fidelity for $0 < \gamma < 0.46$. Finally, if we optimize both, encoder and decoder, with the seesaw iteration, the channel fidelity of the resulting code (E, D) has a strict improvement over the four bit code for $0 < \gamma < 1$.

Second, we look at the five qubit case, that is, the noise is $T_\gamma^{\otimes 5}$. Again, we compare the channel fidelity of different quantum codes (E, D) to the channel fidelity of the single use of the amplitude damping channel $f_c(T_\gamma)$. The results are shown in Figure 3.13. The five bit stabilizer code [57, 58] improves the channel fidelity for $0 < \gamma < 0.27$, the five bit decoder with optimal encoder for $0 < \gamma < 0.29$, and the five bit encoder with optimal decoder for $0 < \gamma < 0.47$. As in the four qubit case, the seesaw iteration results in a code that has a strict improvement over the other codes (including no coding) for $0 < \gamma < 1$.

While this thesis was finalized, Fletcher, Shor, and Win [69] also did the decoder optimizations in Figure 3.12 and 3.13. Their results coincide with the above optimal decoder curves within pixel resolution. Please also note the comment [2] by Werner, Audenaert and myself.

If we compare both cases, $T_\gamma^{\otimes 4}$ and $T_\gamma^{\otimes 5}$, we see that for the optimal codes found by the seesaw iteration, the fidelity gain is below one percent for $0 < \gamma < 0.2$. Given the experimental effort for the additional qubit and that, unlike the assumption of

[19] This has been calculated from the 192 Kraus operators of $E_4 T_\gamma^{\otimes 4} D_{4,\gamma}$ and equation (3.5), $f_c(T) = 1/d^2 \sum_\alpha |\operatorname{tr}(t_\alpha)|^2$.

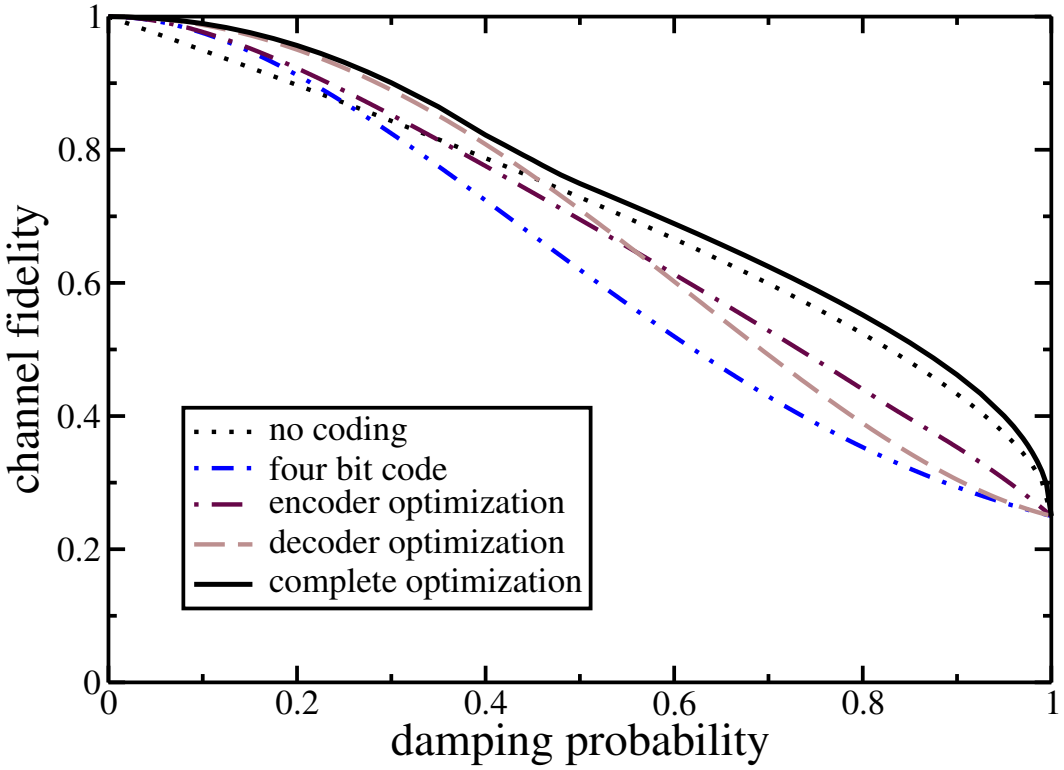

Figure 3.12: Comparison of error correction schemes for the 4-fold tensor product of the amplitude damping channel. The no coding curve displays the function $\gamma \mapsto f_c(T_\gamma)$, the for bit code curve the function $\gamma \mapsto f_c(E_4 T_\gamma^{\otimes 4} D_{4,\gamma})$ with the quantum code by Leung *et al.* [14]. Encoder and decoder optimization takes $D_{4,\gamma}$ and E_4 as decoder and encoder, respectively. The curve of the complete optimization is the result of the seesaw iteration.

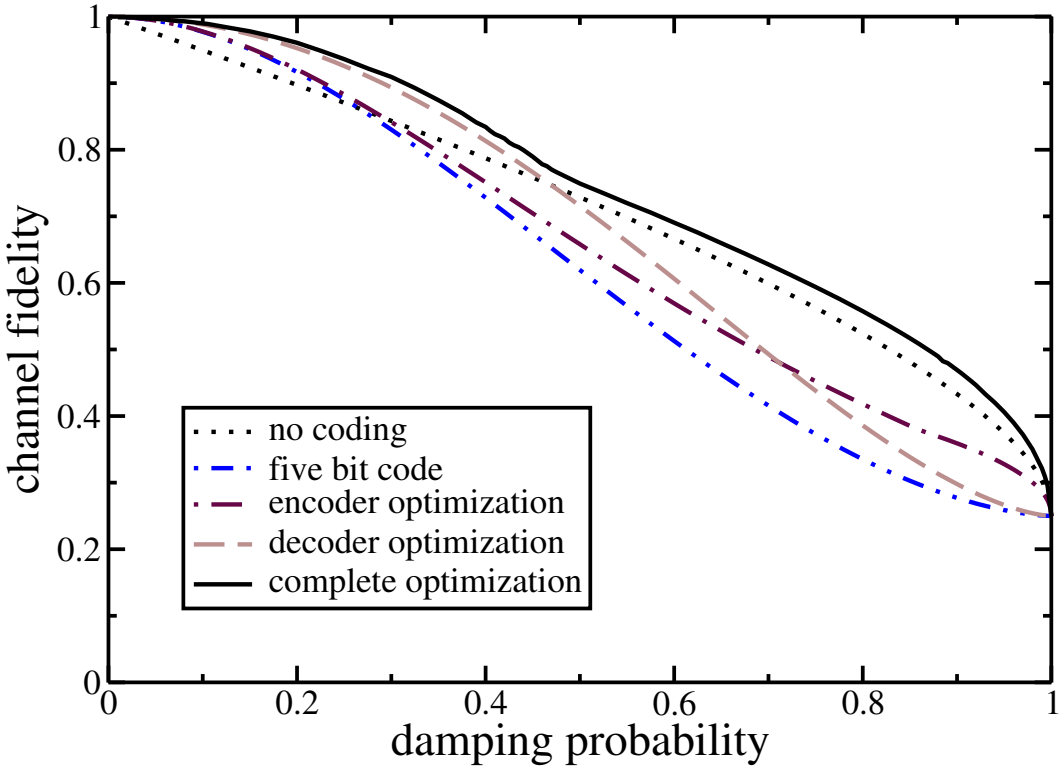

Figure 3.13: Comparison of error correction schemes for the 5-fold tensor product of the amplitude damping channel. The no coding curve displays the function $\gamma \mapsto f_c(T_\gamma)$, the five bit code curve the function $\gamma \mapsto f_c(E_5 T_\gamma^{\otimes 5} D_5)$ with the five bit code [57, 58]. Encoder and decoder optimization takes D_5 and E_5 as decoder and encoder, respectively. The curve of the complete optimization is the result of the seesaw iteration.

the iteration, a quantum circuit can usually not be done with perfect fidelity, it is doubtful that one can increase the fidelity with an additional qubit in the experiment.

Third, we look at the case with the three-fold tensor product $T_\gamma^{\otimes 3}$. We will now compare the optimizations with $f_c(T_\gamma)$ and the above results. One may also compare the results with the majority vote code (E_3, D_3) from equations (3.100) and (3.101). The channel fidelity of the majority vote code for the amplitude damping noise $T_\gamma^{\otimes 3}$ is

$$\begin{aligned} f_c(E_3 T_\gamma^{\otimes 3} D_3) &= \frac{1}{4}\left(\left(1+(1-\gamma)^{3/2}\right)^2 + 3\gamma\left(\gamma-1\right)^2\right) \\ &= \frac{1}{2}\left(1+\sqrt{1-\gamma}-\gamma\sqrt{1-\gamma}-\frac{3}{2}\gamma^2+\gamma^3.\right) \end{aligned}$$

If we compare this to the channel fidelity of the single use of amplitude damping channel as computed in equation (3.105),

$$f_c(T_\gamma) = \left(\frac{1+\sqrt{1-\gamma}}{2}\right)^2 = \frac{1}{2}\left(1+\sqrt{1-\gamma}-\frac{1}{2}\gamma\right),$$

we see that the majority vote code never increases the channel fidelity. This is even true if we optimize the encoder or decoder only, while fixing the other channel to the majority vote version. In particular, the majority vote encoder is already optimal for the majority vote decoder for this noise. The result of the seesaw optimization is shown in Figure 3.14. We see a strict improvement over no coding for $0 < \gamma < 1$. Furthermore, we get an improvement over the four bit code for $\gamma > 0.14$. Figure 3.14 also shows that the seesaw iteration gets frequently stuck for $0.25 < \gamma < 0.4$. Either there is a local optimal solution, or this is due to a flat channel fidelity region.

For the even smaller cases $\max_{E,D} f_c(ET_\gamma D)$ and $\max_{E,D} f_c(ET_\gamma^{\otimes 2} D)$ there is no fidelity improvement over no coding.

In summary, the codes found by the seesaw iteration always outperformed the known error correction codes. There is a strict improvement of the channel fidelity for $0 < \gamma < 1$ in all cases. Even more, for damping parameters $\gamma > 0.14$, the optimal code found by the seesaw iteration for $T_\gamma^{\otimes 3}$ is superior to the four bit code [14] for $T_\gamma^{\otimes 4}$. On the other hand, this also suggests that, although not optimal, the four bit code is better than any three qubit code for small damping probabilities. Finally, Figure 3.14 shows that the seesaw iteration can result in a code that is not globally optimal.

3.3.4 Tough Error Models

We will now use the iteration to study the ability to correct a noisy quantum channel, given only the number of independent Kraus operators of the channel. Consider a noisy channel T on a n-dimensional system, which requires at most k Kraus

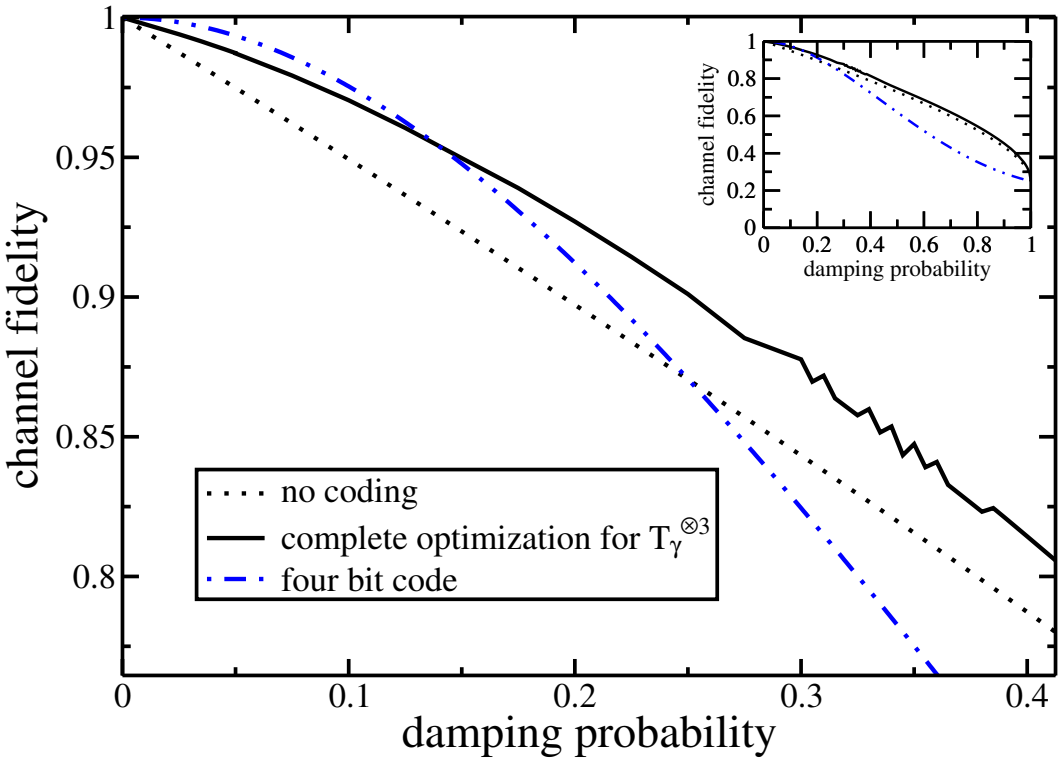

Figure 3.14: Comparison of the channel fidelity of the single use of the amplitude damping channel T_γ without coding with the optimal codes found by the seesaw iteration for the three-fold tensor product of the amplitude damping channel $T_\gamma^{\otimes 3}$, and the four bit code [14] applied to the four-fold tensor product of the amplitude damping channel $T_\gamma^{\otimes 4}$.

operators. If k is sufficiently small, one can find a perfect error correcting code (E, D) for the transmission of a c-dimensional system,

$$f_c(ETD) = f_c(\mathrm{id}_c) = 1.$$

We will obtain analytic and numerical results on the triples (c, n, k) making such correction possible without further information on the error operators.

In terms of the Stinespring representation, the restriction to k Kraus operators means that the dilation space is at most k-dimensional. So in this sense, by restricting k, we set an upper bound on the dimension of the environment the noise interacts with. This is where the intuition comes from: If the noise does not interact with the environment, i. e., $k = 1$, the evolution of the system is unitary and therefore completely reversible. Thus, if the noise only interacts with a small part of the environment, $k \ll n$, we expect it to be partly reversible, with a tradeoff between k and the dimension c of the reversible subspace. This brings us to the following questions.

3.3.1 Problem. Consider noisy channels of the form $T\colon \mathcal{B}(\mathcal{H}) \to \mathcal{B}(\mathcal{H})$, $\dim \mathcal{H} = n$, with at most k Kraus operators, $T(A) = \sum_{\alpha=1}^{k} t_\alpha^* A\ t_\alpha$. Furthermore, consider all codes (E, D) with $E\colon \mathcal{B}(\mathcal{H}) \to \mathcal{B}(\mathcal{K})$, $D\colon \mathcal{B}(\mathcal{K}) \to \mathcal{B}(\mathcal{H})$, $dim\mathcal{K} = c$. We then ask:

For which triples (c, n, k) are all channels correctable?

That is, there exist a perfect error correction scheme with code dimension c for every channel T that maps between bounded operators on a n dimensional Hilbert space and has k Kraus operators.

Certainly, some combinations of error operators are harder to correct than other. In this regard, we will call the boundary cases of (c, n, k) *tough error models*. Also, one can restrict oneself to consider only linearly independent Kraus operators. For linearly dependent Kraus operators, one can find a Kraus representation with less operators, hence if such a channel is correctable, it is already correctable using fewer Kraus operators. Lemma 2.2.11 on page 15 implies that we have at most n^2 linear independent Kraus operators for any channel T.

Tough error models are interesting for both, small and large parameters of the dimensions c and n. Small parameters can be interesting for actual implementation, where the asymptotic behavior provides bounds on the feasibility of quantum computing. In this chapter, we will only consider error correcting codes with isometric encoding (see Definition 3.1.3 on page 25). Therefore, we can focus on the Knill-Laflamme condition (3.2) from Theorem 3.1.4 on page 26, as it is necessary and sufficient for the existence of a quantum code with isometric encoder: There exists a perfect error correcting code with isometric encoder $E(A) = v^* A v$ that corrects a set of operators $\mathcal{E}$ if and only if for all $e, f \in \mathcal{E}$

$$v^* e^* f v = \omega(e^* f)\mathbb{1}, \tag{3.109}$$

with complex numbers $\omega(e^* f)$.

The Knill-Laflamme condition (3.109) does not take into account that the error operators e, f in question belong to a Kraus decomposition of the noise channel and hence fulfill a channel constraint. But taking the channel constraint into account does not change the behavior of bounds on (c, n, k). If (c, n, k) is correctable for k Kraus operators t_α without channel constraint, then (c, n, k) is, of course, also correctable for k operators t_α with channel constraint $\sum_\alpha t_\alpha^* t_\alpha = \mathbb{1}$. But this implies that $(c, n, k-1)$ is correctable for $k-1$ operators without channel constraint by setting

$$t_k = \sqrt{\mathbb{1} - \sum_{\alpha=1}^{k-1} t_\alpha^* t_\alpha}$$

and hence the asymptotic behavior of bounds is not affected.

3.3.4.1 Bounds

A simple upper bound on k for the problem is given by entanglement breaking channels [76]. Entanglement breaking channels destroy the initial entanglement of a state with another system. That is, the channel T_* is entanglement breaking, if $\mathrm{id} \otimes T_*(\rho)$ is separable for all ρ. Since entanglement can't be increased by local operations alone, we cannot find a code (E, D) such that ETD is the identity. Entanglement breaking channels can be written as [76]

$$T(A) = \sum_{j=1}^{n} |\Psi_j\rangle\langle\Phi_j|A\Phi_j\rangle\langle\Psi_j|,$$

with some Hilbert space vectors $|\Psi_j\rangle, |\Phi_j\rangle$. That is, the number of Kraus operators k is at most n, the dimension of the underlying Hilbert space $\mathcal{H}$. This implies the upper bound $k \leq n$ for the admissible triples (c, n, k).

We will now develop a lower bound on the code dimension c. The bound is based on the idea to realize the Knill-Laflamme condition (3.109) in two steps: First, we realize the independency of $\omega(e^* f)$ from the isometry v. Then, we establish the orthogonality of the codewords, that is, the orthogonality of the columns of v, which realizes the $\mathbb{1}$ in equation (3.109). With this strategy, we obtain the following result.

3.3.2 Proposition. *Let n be the dimension of the Hilbert space $\mathcal{H}$ of a noisy channel $T: \mathcal{B}(\mathcal{H}) \to \mathcal{B}(\mathcal{H})$ with k Kraus operators. Then one can find a perfect error correcting code with isometric encoder for code dimension c, such that*

$$c \geq \left\lfloor \frac{n}{2^{k^2}} \right\rfloor. \tag{3.110}$$

The proof of this bound is based on the following Lemma.

3.3.3 Lemma. *Let $A \in \mathcal{B}(\mathbb{C}^n)$ be a hermitian operator. Then there exists a scalar ω and at least $n/2$ vectors $q_i \in \mathbb{C}^n$ such that*

$$\langle q_i | A q_j \rangle = \omega \delta_{ij}.$$

Proof. Let $A|\Phi_i\rangle = \lambda_i |\Phi_i\rangle$, $i = 1, \dots n$, be the eigendecomposition of A. Let ω be the median of the eigenvalues. Now successively take the largest and smallest eigenvalue and define

$$q_i(t) := \cos(t)|\Phi_i\rangle + \sin(t)|\Phi_{n-i}\rangle.$$

By intermediate value theorem there exists a t such that

$$\langle q_i(t) | A q_i(t) \rangle = \omega.$$

■

With this Lemma, we can construct an isometry v that satisfies the Knill-Laflamme condition (3.109).

Proof of Proposition 3.3.2. Given a noisy channel characterized by the dimension n with k Kraus operators, we choose a hermitian basis A_j for the linear span of all operators of the form $e^* f$, where e and f are the Kraus operators of the noisy channel. As there are k^2 combinations $e^* f$, the hermitian basis consists of at most k^2 operators. Note that $e^* f$ are not generally hermitian, so we need to decompose them into hermitian and anti-hermitian parts. But for every combination $e^* f$, the combination $f^* e$ is also in that span, i. e., the set of operators $e^* f$ is closed under the involution, and hence there is no factor two in the estimation of the number of hermitian basis elements.

Now we successively apply the above Lemma to the basis operators A_j, restricted to the subspace spanned by the q_i of the previous step. As this halves the remaining dimension in every step and A_1 has n eigenvectors, the final subspace is $\lfloor n/2^{k^2} \rfloor$-dimensional. For the vectors $|q_i\rangle$ of the last step, we have

$$\langle q_i | e^* f | q_j \rangle = \omega(e^* f)\delta_{ij},$$

for all combinations $e^* f$. This leads us to an encoding isometry $v = \sum_i |q_i\rangle\langle i|$ satisfying (3.109), where $|i\rangle$ is a basis of a $\lfloor n/2^{k^2} \rfloor$-dimensional Hilbert space. ■

For small numbers of Kraus operator the bound (3.110) yields:

k	1	2	3	4	5
$n \leq$	$2c$	$16c$	$512c$	$65536c$	$33554432c$

For a single Kraus operator the channel is unitary and hence invertible, leading to a code with $c = n$. In this case the bound gives $c \geq \lfloor n/2 \rfloor$, and therefore the bound isn't tight. The first nontrivial case $k = 2$ gives $c \geq n/16$, which is the same ratio as for the five bit stabilizer code [57, 58] that corrects 16 specific Kraus operators. On the other hand, correcting 16 generic Kraus operators according to the bound strategy already requires $n = 2^{256}c$.

Knill, Laflamme and Viola [71] obtained a lower bound by realizing the Knill-Laflamme condition (3.109) the other way round: They first produce the orthogonality $\mathbb{1}$, and then the independency of $\omega(e^*f)$ from the isometry v. With this approach they got the following bound on the code dimension.

3.3.4 Proposition (Knill, Laflamme, Viola)**.** *Let n be the dimension of the Hilbert space $\mathcal{H}$ of a noisy channel $T\colon \mathcal{B}(\mathcal{H}) \to \mathcal{B}(\mathcal{H})$ with k Kraus operators. Then one can find a perfect error correcting code with isometric encoder for code dimension c, such that*

$$c \geq \left\lceil \frac{n}{k^2} \right\rceil \frac{1}{k^2+1}. \tag{3.111}$$

They obtain this bound as follows. First, they use a classical code with $c \geq \lceil \frac{n}{k^2} \rceil$ such that

$$\langle c_i | e^* f c_j \rangle = \omega_i(e^*f)\delta_{ij}$$

holds for the classical codewords $|c_i\rangle$. Note that ω still depends on i, so this is not already a quantum code. Then, they use convex combinations of the classical codewords as quantum codewords

$$|q_i\rangle = \sum_{j \in I_i} \sqrt{\beta_{ij}} |c_j\rangle \quad \text{with} \sum_{j \in I_i} \beta_{ij} = 1,$$

where β_{ij} are the coefficients and I_i are certain index sets about to be defined. Doing so one gets

$$\langle q_i | e^* f q_j \rangle = \sum_{j \in I_i} \beta_{ij} \omega_j(e^*f)\delta_{ij}. \tag{3.112}$$

So the coefficients β_{ij} and index sets I_i have to be chosen such that $\sum_{j \in I_i} \beta_{ij}\omega_j(e^*f)$ is independent of i. This can be done according to Radon's Theorem [77].

3.3.5 Theorem (Radon)**.** *Any set of $(c(m+1)-m)$ points in a real m-dimensional vector space can be divided into c sets, whose convex hulls have non-empty intersection.*

Note that Radon's Theorem is tight, i. e., there exists counter examples for the division into $c+1$ sets. To apply the Theorem, we take the ω_j as k^2 dimensional real vectors. The real dimension is k^2 since for every combination e^*f we also have the combination f^*e. As the dimension of the classical code is $\geq \lceil \frac{n}{k^2} \rceil$, the Theorem

tells us that we can find c sets I_i whose convex hulls have a common point ω in their intersection, as long as

$$\left\lceil \frac{n}{k^2} \right\rceil \leq c(k^2+1) - k^2.$$

In that case, we take β_{ij} in equation (3.112) to be the convex weights that correspond to the common point ω in the intersection, i. e., we have

$$\sum_{j \in I_i} \beta_{ij} \omega_j(e^* f) = \omega(e^* f)$$

independent of i. This leads us to the isometry $v = \sum_{i=1}^{c} |q_i\rangle\langle i|$ that satisfies the Knill-Laflamme condition (3.109). Thus a perfect error correcting code exists for

$$\left\lceil \frac{n}{k^2} \right\rceil \frac{1}{k^2+1} \leq c - \frac{k^2}{k^2+1},$$

which concludes the proof of Prop. 3.3.4.

For small numbers of Kraus operator the bound (3.111) yields:

k	1	2	3	4	5
$n \leq$	$2c$	$20c$	$90c$	$272c$	$650c$

The case of a single Kraus operators shows that this bound is not tight, also. For two Kraus operators, $k = 2$, we obtain that $c \geq n/20$, which is slightly worse than the $c \geq n/16$ we got from the former bound (3.110). For $k \geq 3$ this bound is always better. Correcting 16 arbitrary Kraus operators according to this strategy leads to $n = 65792c$. This is still a very large overhead compared to the correction of the 16 localized Pauli errors that can be corrected using the five bit code. For this noise, we have $n = 2^5 = 2^4 c$ using the five bit code. Furthermore, observe that the five bit code also improves the channel fidelity for the five-fold tensor product of the depolarizing channel with depolarization probability $p < 1 - \sqrt{6}/3 \approx 0.18$ as we have shown in section 3.3.2. That is, although it is not a perfect error correcting code in this case, it improves the fidelity in the presence of 4^5 Kraus operators of the noise.

Since isometric encoding is sufficient in terms of quantum capacity, equation (3.111) also implies a bound on the quantum capacity. If we set $n = D^N$ and $k = K^N$, then, for large N, we get

$$c \geq \frac{1}{2} \left(\frac{D}{K^4} \right)^N,$$

which leads to a bound on the quantum capacity Q,

$$Q = \lim_{N \to \infty} \frac{\log c}{N} \geq \log D - 4 \log K.$$

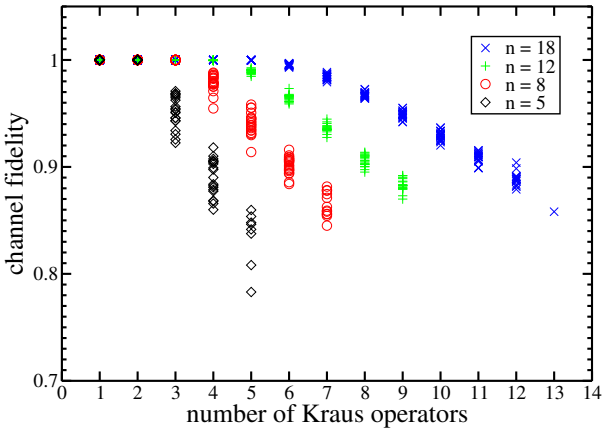

Figure 3.15: Iteration results for the encoding of qubits, $c = 2$, into n-dimensional Hilbert spaces subject to random noisy channels $T: \mathcal{B}(\mathcal{H}) \to \mathcal{B}(\mathcal{H})$, $\dim \mathcal{H} = n$, with various number of Kraus operators.

3.3.4.2 Numerical Results

We will now numerically investigate the triples (c, n, k) in the qubit case $c = 2$. For a given noisy channels T, we will use the seesaw iteration with semidefinite programming for the single iteration steps to find the optimal code (E, D). As the channel fidelity $f(ETD)$ is one if and only if ETD is the ideal channel, we can show, up to a small threshold, whether a given channel can be corrected perfectly. Every channel that we cannot correct perfectly then suggests an upper bound on the triples (c, n, k). However, bear in mind that the seesaw iteration does not guarantee that the code found is indeed a global optimum, so the bound could be too pessimistic.

At first, we will use random channels as noise. The random channels are created by choosing real and imaginary part of a matrix uniformly distributed between 0 and 1 and using Gram-Schmidt orthogonalization to create a Stinespring isometry for the noisy channel out of it. The number k of Kraus operators is determined by the dimension chosen for the dilation space. The following table lists the number of Kraus operators for which the channel fidelity of the optimal code found by the seesaw iteration is below 99 percent, $f(ETD) < 0.99$.

n	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
k	2	3	3	3	4	4	4	5	5	5	6	6	6	6	7	7	7	7

Figure 3.15 shows results of the seesaw iteration for some choices of n. One can see that the combinations $(2, 5, 3)$, $(2, 8, 4)$, $(2, 12, 5)$ and $(2, 18, 7)$ are not correctable with $f_c > 0.99$ and form a numerical upper bound on the (c, n, k). If we compare this result with the lower bound (3.111), we see that lower bound yields $n \leq 180$ already for the case $k = 3$. This encourages to search for both, tougher noise and

better bounds.

We will now look at the correction of $SU(2)$ covariant channels. The motivation for this comes from the following observation.

3.3.6 Proposition. *No isometric encoding v that perfectly corrects localized errors can be covariant in the sense that*

$$vu = u^{\otimes n}v$$

for all unitary u.

Proof. Let v be a covariant encoding that corrects localized errors. So the encoding channel is $E(A) = v^*Av$ with $vu = u^{\otimes n}v$ for all unitary u. This implies that for any hermitian operator X,

$$v^* \left(e^{-itX}\right)^{\otimes n} v = e^{-itX}.$$

Differentiating at $t = 0$ leads to

$$i\frac{\mathrm{d}}{\mathrm{d}t}v^*(e^{itX})^{\otimes n}v\Big|_{t=0} = v^* \left(\sum_{j=1}^{n} X_j \otimes \mathbb{1}_{n\setminus j}\right) v = X, \tag{3.113}$$

where j and $n \setminus j$ are the index sets of the corresponding tensor factors. On the other hand, as v corrects localized errors, it satisfies the Knill-Laflamme condition (3.109). Thus we have

$$\sum_{j=1}^{n} v^* X_j \otimes \mathbb{1}_{n\setminus j} v = \sum_{j=1}^{n} \omega(X_j \otimes \mathbb{1}_{n\setminus j})\mathbb{1}. \tag{3.114}$$

In order to satisfy both equations, (3.113) and (3.114), X has to be a multiple of $\mathbb{1}$. This means that v can't correct single localized qubit errors, and in particular, no nontrivial distance based stabilizer encoding can be covariant. ■

More general, error operators that are generators of the symmetry of the encoding isometry can't be corrected unless the corresponding Lie-algebra is abelian in which case v is a classical code.

This suggest to look at the case of covariant noise. Let t be an $SU(2)$ covariant Stinespring isometry of the channel T, i. e.,

$$T(A) = t^*(A \otimes \mathbb{1})t,$$
$$tD^j = (D^j \otimes D^s)t,$$

where D^j and D^s are representations of $SU(2)$. Here D^s corresponds to the dilation space with dimension $(2s+1)$. The reasonable choices of s are limited by the maximal minimal dilation dimension $n^2 = (2j+1)^2$,

$$(2s+1) \leq (2j+1)^2.$$

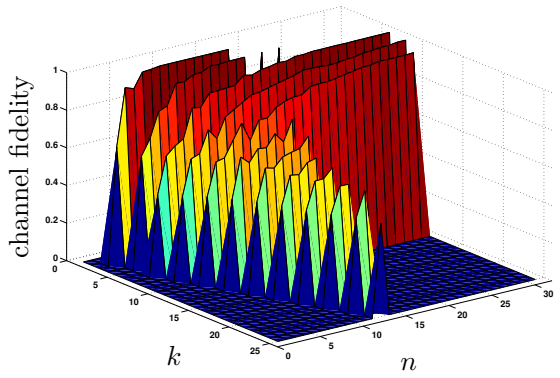

Figure 3.16: Iteration results of the encoding of qubits, $c = 2$, for $SU(2)$ covariant noisy channels $T: \mathcal{B}(\mathcal{H}) \to \mathcal{B}(\mathcal{H})$, $\dim \mathcal{H} = n$, with various number k of Kraus operators.

The dimension of D^s also implies that the number of Kraus operators is odd. The isometry t is given by Clebsch-Gordon coefficients. From this we get the additional constraint

$$|j - s| \leq j \leq j + s.$$

Again, we use the seesaw iteration to investigate the qubit case $c = 2$. The results are shown in the Figures 3.16 and 3.17. In Figure 3.16 we see that the fidelity is the better, the fewer Kraus operators the noise has and the larger the dimension of the noisy system is. Furthermore, we see that for a fixed number of Kraus operators k, the iterated codes sometimes decrease the channel fidelity with larger n. However, as the seesaw iteration was stopped if the gain in fidelity was below 10^{-8}, and several runs have been made for these parameters, these cases seem to have a flat fidelity region around the optimal value. This is interesting, because it shows the limitations of a numerical approach to quantum error correction. It can also be seen in the following table, that lists the number of Kraus operators for which the channel fidelity of the optimal code found by the seesaw iteration is below 99 percent, $f(ETD) < 0.99$.

n	2	...	6	7	...	11	12	13	14	...	21	22	...	29	30	31
k	3	...	3	5	...	5	7	5	7	...	7	9	...	9	11	11

While comparing these result with the results for random noise, note that, as shown above, the number of Kraus operators of a $SU(2)$ covariant channel is always odd. Figure 3.17 shows that the (c, n, k) combinations $(2, 5, 3)$, $(2, 12, 5)$, and $(2, 19, 7)$ impose the same bound as in the case of random noise. It also shows that $(2, 27, 9)$ is not correctable. In this sense, covariant noise is as tough as random noise.

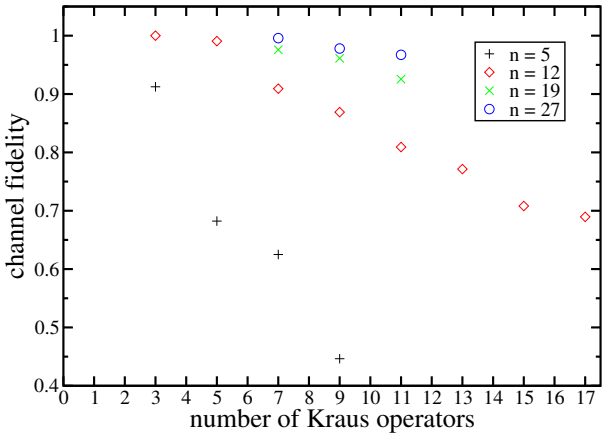

Figure 3.17: Iteration results of the encoding of qubits,$c = 2$, for $SU(2)$ covariant noisy channels $T\colon \mathcal{B}(\mathcal{H}) \to \mathcal{B}(\mathcal{H})$, $\dim \mathcal{H} = n$, for selected values of n.

3.3.4.3 Conclusion

We have proven another lower bound for tough error models that realizes the Knill-Laflamme condition by spectra flattening and filtering instead of orthogonalization and Radon's theorem, as done for the bound by Knill, Laflamme, and Viola. Neither bound is tight, but the new bound is not based on the usage of classical codes and provides better results in the case $k = 3$. Furthermore, we have shown that isometric encoding with a covariant isometry does not allow to correct local errors. In particular, this means that no distance based stabilizer can have a covariant encoding isometry. Additionally, we used the seesaw iteration to calculate upper bounds on the triples (c, n, k) using random and $SU(2)$-covariant noise. As there is a fairly large gap between upper and lower bounds, this suggest that there exist tougher error models or better bounds.

3.3.5 Classical Feedback

In contrast to the coding schemes above, where we encode a smaller quantum system into a larger one, we will now look at the case of recovery only. That is, the dimension of the quantum system stays the same and there is no encoding operation. However, we will allow recovery operations during and after the time evolution conditioned on the outcomes of measurements on the environment. This setup is known as classical feedback in discrete time [78]. Using a seesaw iteration with the subchannel power iteration, we will show that the optimal feedback strategies for qubits are not optimal for higher level systems.

Let $T\colon \mathcal{B}(\mathcal{H}) \to \mathcal{B}(\mathcal{H})$, $\dim \mathcal{H} = d$, be a noisy channel that is given by n Markovian

steps,

$$T = T^{(1)} \circ \cdots \circ T^{(n)},$$

with n possibly different channels $T^{(k)}$. We allow correction operations between any two subsequent steps that can be conditioned on all measurement results obtained so far. From the Stinespring representation of $T^{(k)}$ it is clear that we can associate a Kraus representation $T^{(k)}(A) = t^*_{\alpha_k} A t_{\alpha_k}$ with any measurement on the environment. If we take the environment large enough, it suffices to consider projective measurements[20], but any measurement of a basis on the dilation space in the Stinespring representation corresponds to a Kraus decomposition of the channel. So the most general correction scheme is

$$T_{corr,*}(\rho) = \sum_{\alpha_1,\ldots\alpha_n} R_*^{(x_1,\ldots x_n)} \left(t^{(n)}_{\alpha_n} \ldots R_*^{(x_1)} (t^{(1)}_{\alpha_1} \rho \, t^{(1)*}_{\alpha_1}) \ldots t^{(n)*}_{\alpha_n} \right).$$

The main idea can be seen from the example of the depolarizing channel. From (3.7) on page 28 we know that the depolarizing channel has a Kraus decomposition, where each Kraus operator is proportional to a unitary operator. If we do a measurement on the environment such that we know which Kraus operator was applied, we can simply revert the unitary evolution and end up with the ideal channel. This works independently of the depolarization probability. In particular, we can perfectly recover from the completely depolarizing channel, which has zero quantum capacity. In general, in the single step situation, i. e., $n = 1$, Gregoratti and Werner obtain (Prop. 4 [78])

$$f_c(T_{corr}) \leq 1/d^2 \sum_\alpha (\operatorname{tr} |t_\alpha|)^2$$

via the Cauchy-Schwarz inequality. So the optimal recovery is given by the polar isometries (unitaries) of the Kraus operators t_α.

In the multi-step case, they look at different strategies as depicted in Figure 3.18. In (a), the most general case is shown. The two special cases (b) and (c) correspond to the strategy to condition the recovery only on the latest measurement result, respectively, to do a single correction at the end. Gregoratti and Werner [78] showed that both special cases are equivalent and optimal for qubit systems.

For $d = 3$, we will use a seesaw iteration combined with the subchannel power iteration to show that neither case, (b) or (c), is generally optimal. As a counterexample, consider a two-step channel, $T = T^{(2)} \circ T^{(1)}$, with $T^{(1)} = T^{(2)}$ and $T^{(1)}$

[20] See chapter 2.2.8 of [35] how to turn a POVM into a PVM on a larger system.

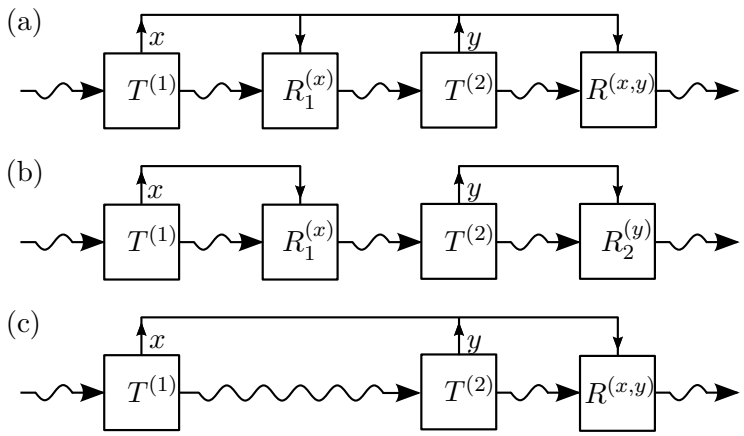

Figure 3.18: Different strategies for error correction based on classical feedback for the multi-step channel $T = T^{(2)} \circ T^{(1)}$ [78]. In the most general case (a), the final correction operations R depends on all information gathered so far. This includes the special cases (b) and (c). In (b), only the last measurement result is used for the correction. In (c), there is a single correction operation at the end, depending on all measurement results obtained.

given by the Kraus operators

$$t_1 = \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} \sqrt{1/6} & 0 & 0 \\ 0 & \sqrt{1/3} & 0 \\ 0 & 0 & \sqrt{1/2} \end{pmatrix},$$
$$t_2 = \begin{pmatrix} \sqrt{5/6} & 0 & 0 \\ 0 & \sqrt{2/3} & 0 \\ 0 & 0 & \sqrt{1/2} \end{pmatrix}.$$

The operator t_1 is already written in polar decomposition. The fidelities given by the strategies (b) and (c) are

$$f_c^{(b)}(T_{corr}) = \frac{1}{d^2} \sum_{\alpha_1,\alpha_2} \left(\mathrm{tr}\left|t_{\alpha_2}|t_{\alpha_1}|\right|\right)^2 \approx 0.9570,$$
$$f_c^{(c)}(T_{corr}) = \frac{1}{d^2} \sum_{\alpha_1,\alpha_2} \left(\mathrm{tr}\left|t_{\alpha_2}t_{\alpha_1}\right|\right)^2 \approx 0.9556.$$

We will compare these channel fidelities with the numerical result of the seesaw iteration for

$$\max_{Q_{\alpha,\beta},R_\alpha} \sum_{\alpha,\beta} f_c\left(T_\alpha R_\alpha T_\beta Q_{\alpha,\beta}\right), \tag{3.115}$$

which corresponds to the most general case (a) in Figure 3.18. Here, we already used the linearity of f_c, so the optimization can be done separately for each α or

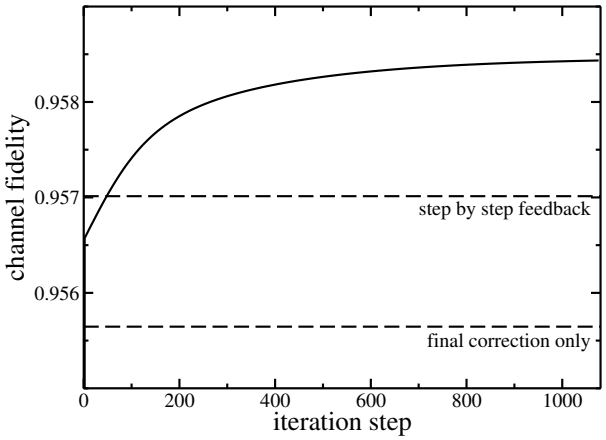

Figure 3.19: Result of the seesaw iteration (3.115) for optimal feedback for the channel $T = T^{(2)} \circ T^{(1)}$ defined in the text.

(α, β). From equation (3.33) on page 35, we see that the corresponding Jamiolkowsky operators can be calculated via

$$\langle\!\langle \beta | \widetilde{F} | \alpha \rangle\!\rangle = f(|\alpha\rangle\!\rangle\langle\!\langle\beta|).$$

The result of the iteration is shown in Figure 3.19. The fidelity obtained by the seesaw iteration is approximately 0.9584. We see that neither a step-by-step correction, nor a single correction at the end is optimal. This is interesting, because it implies that the optimal correction requires some foresight in the intermediate correction R_α. Since $T^{(1)} = T^{(2)}$ and $\big||t_2|t_2|\big| = |t_2||t_1|$, optimal correction based on the measurement results so far is equivalent to optimal step-by-step correction (b). Gregoratti and Werner also give an example where step-by-step correction (b) is worse than final correction (c). Also in this case, both are inferior to the iteration result.

3.4 Conclusion

Error correction can be formulated as an optimization problem with the channel fidelity as linear objective, and the encoder channel E and decoder channel D as entities to optimize. The seesaw iteration solves the joint optimization over the pair of channels (E, D) numerically by alternatingly optimizing the encoder with fixed decoder and optimizing the decoder with fixed encoder, unless the gain in fidelity is below some threshold. This results in a local optimal pair of encoder and decoder channel for the given noise.

The involved optimizations of an objective functional $f(S)$ over all channels S for $S = E$ and $S = D$ are both convex optimization problems. They can be solved

via semidefinite programming. Furthermore, we developed iterative algorithms for the optimization of $f(S)$: the subchannel power iteration and the channel power iteration. These iterations monotonically increase the objective functional and every global optimal solution of the channel power iteration is a stable fixed point. Moreover, they have a high numerical stability since errors do not accumulate over iteration steps. Compared to a standard semidefinite programming approach, they require less computational effort and allow to restrict the number of independent Kraus operators. For example it is possible to restrict E to isometric encodings, which is optimal in an asymptotic sense and often also turns out to be optimal in low dimensions.

In general, the optimization scheme makes neither prior assumptions about the structure of coding and decoding operation nor does it make prior assumptions about the noise. In this sense, the seesaw iteration corresponds to an ab-initio-approach to quantum error correction. The iterated encoders and decoders adapt to the given noise, they utilize symmetries of the global dynamics on demand and even find the special form of Knill-Laflamme type of codes where they are optimal. In the case where no perfect error correcting code exists, the code finds approximate error correction schemes.

In particular, the seesaw iteration finds a perfect error correction scheme in the presence of a noiseless subsystem in the example of the three qubit collective rotation channel. Furthermore, for the three-fold tensor product of the bit-flip channel, the only quantum error with a classical analogue, a code that is unitarily equivalent to the classical majority vote code is found.

For the five-fold tensor product of the depolarizing channel for damping parameter less or equal one, we could not find any better code than choosing either the five bit stabilizer code or no coding. This is remarkable, as it suggests that the five bit code is already optimal in a worst case error scenario, although it is designed to correct rare errors only. In contrast, new codes are found for damping parameters greater one, and for the n-fold tensor product of the amplitude damping channel for $n = 3, 4, 5$. These codes are even better than the approximate error correcting four bit code by Leung *et al.* in the case $n = 4$, and better than the five bit code for $n = 5$. The novelty of the codes is deduced from the fact that they lead to superior fidelities compared to the known codes, even in the cases where the encoder or decoder of the known codes were optimized to the specific noise. Like the four bit code, the new code found for $n = 4$ violates the singleton bound for distance based perfect quantum error correction.

The seesaw iteration was used to provide upper bounds on tough error models, that is, on the ability to perfectly correct noisy channels where only the number of independent Kraus operators is fixed. These results were compared to a known and a newly developed lower bound. The large difference between lower and upper

bounds suggests that there exist better bounds or tougher errors.

For error correction using classical feedback in discrete time we used a seesaw iteration to show that a correction strategy, that is proven to be optimal in the qubit case, is not optimal for higher dimensional systems. This is an example for the usage of a numerical method to test conjectures and either falsify them or build up further confidence.

In summary, the subchannel power iteration provides a reliable and efficient algorithm for the optimization of any linear objective over the set of channels. Combined with the seesaw iteration, it makes ab-initio error correcting available to test the performance of perfect as well as approximate error correction schemes.

Chapter 4

Postprocessing Tomography Data

An experimental realization of a quantum system usually differs more or less from its design that is based on a theoretical model. Typically, the model does not account for the given small imperfections in the laboratory. For the analysis and verification of the quantum system, it has to be estimated using the accessible measurements. This identification of an experimentally implemented quantum system via measurements is called tomography. Quantum state tomography refers to the identification of quantum states, and quantum channel tomography or quantum process tomography refers to the identification of quantum channels.

The tomography of a quantum system yields a sequence of measuring results for measurements given by the tomography strategy. As a consequence of the statistical and systematic errors of the measurement, a direct interpretation of the data has the problem that physical constraints are often not met. In particular, the positivity condition for quantum states and the unitarity condition for quantum gates are typically violated [79]. In this sense, the measured data suggests some unphysical behavior. To regain a physical meaning, one searches a model quantum system that complies best with the measured data. The assignment of such a model to the data is done via a fitting algorithm. A common method for this is the maximum-likelihood estimation [17, 16] which will be explained in the following section. In [18], Kosut *et al.* show that such a fitting of a state or channel to the measurement results can be formulated as a conic program. Thus, this problem is efficiently solvable numerically.

However, already in the case of state tomography, given a d dimensional Hilbert space, a full tomography requires the measurement of $d^2 - 1$ degrees of freedom (real diagonal plus complex upper triangle minus normalization). This exponential growth of the required number of measurements with increasing number of qubits for a multi-qubit state raises the question, whether it is possible to gain meaningful insight into

properties of the engineered quantum system with fewer measurements. In this chapter, it is shown how to extend the postprocessing in reference [18] to obtain upper and lower bounds for linear figures of merit for any number of measurement results. As an application of this method, it is shown how to obtain the best-case fidelity and worst-case fidelity to a designated pure state or a designated unitary channel. For the latter, we will once again use the channel fidelity defined in the previous chapter as linear objective. From the analysis of tomography data, it will become apparent that the fidelity measure is highly sensitive to the chosen postprocessing method.

4.1 General Framework

Especially if the theoretical model of the engineered quantum system is a pure state or a unitary channel, conditions on the representing matrix such as positivity are often not met. Pure states and unitary channels are the extreme points of the corresponding probability distributions, and already a slight deviation would lead to "probabilities" greater than one or less than zero. Instead of a direct interpretation of the measurement results, we ask the question:

$$\text{Which quantum system is likely to produce the measured data?} \tag{4.1}$$

This question is answered by maximizing the likelihood function. The likelihood is the probability to get the given measurement results assuming one has a specific system in the laboratory. The system that is assigned to the measurement results is one that maximizes this probability.

Let n_i be the number of times outcome i was obtained in a measurement of $\{m_i\}$, and let $s(m_i)$ be the probability of obtaining outcome i given the quantum system is s. If we assume that the experiments are statistically independent, the probability of obtaining the data n_i is given by

$$\prod_i s(m_i)^{n_i} \tag{4.2}$$

For example, in the case of states, m_i would be an operator of a POVM, $s = \rho$, and $s(m_i) = \mathrm{tr}(\rho m_i)$.

The quantum system that maximizes (4.2) is the one that maximizes the log-likelihood function

$$L(s) := \log \prod_i s(m_i)^{n_i} = \sum_i n_i \log s(m_i). \tag{4.3}$$

As non-negative weighted sum of concave functions, (4.3) is a concave objective and the optimization can be done via a downhill simplex method in the case of

state [15] and channel tomography [16]. Recall that for conic programs, given the duality gap is zero, the dual optimal solution certifies that the primal solution found is indeed a global optimum. Yet, the simplex method does not exploit duality to prove optimality of the result. Hence it is possible to get stuck in a particular flat region of the objective. The problem in applying duality theory to (4.2) or (4.3) is that the objective is not linear in s.

If, however, we assume independent normal distribution of the measured data, the probability that this data is produced by a physical system s becomes proportional to

$$\prod_i e^{-\frac{1}{2}\left(\frac{d_i - s(m_i)}{\sigma_i}\right)^2}. \tag{4.4}$$

Here, d_i is the relative frequency of the result i corresponding to the measurement of $\{m_i\}$ with standard deviation σ_i, and $s(m_i)$ is the probability of the result i, given the quantum system is s. The probability (4.4) can be maximized by solving the minimization

$$\min_s \sum_i \left(\frac{d_i - s(m_i)}{\sigma_i}\right)^2, \tag{4.5}$$

which is known as least squares minimization.

Note that often measurement results are obtained by counting events, e. g., clicks in an avalanche photodiode. In these cases, measurement errors are usually Poisson distributed. According to the central limit theorem, a Poisson distribution can be approximated by a normal distribution, if the expectation number of clicks gets large [80]. However, if this is not the case, this approximation may result in a misleading maximum likelihood estimation. Also note that the normal distribution is not robust against outliers. Furthermore, observe that the standard deviation is $\sigma_i = \sqrt{Np_i}$ in the case of Poisson distributed measurement data, where N is the number of trials and p_i is the probability of getting a click. Thus, when we expect no clicks from the model, i. e., $p_i = 0$, we have $\sigma_i = 0$, and few clicks due to noise in the experiment get an infinite weight in the corresponding least squares term in (4.5), which can greatly reduce the fidelity of the fit.

The main virtue of the formulation (4.5) is that with fixed values of the σ_i and a standard trick of convex optimization about to be mentioned, the objective (4.5) can be written as the optimization of a linear functional with a conic constraint. Thus, it can be solved via conic programming [63, 18]. A conic program is the optimization of a linear objective, where the variable is restricted to lie in the intersection of an affine plane with a convex cone. The trick is that instead of solving (4.5), we minimize an auxiliary variable t, such that

$$t^2 \geq \sum_i \left(\frac{d_i - s(m_i)}{\sigma_i}\right)^2. \tag{4.6}$$

Now (4.5) can be written as conic program of the form

$$\min_x \{\langle c|x\rangle | Ax - b \geq_{\mathcal{C}} 0\} \,, \tag{4.7}$$

where c and x are complex vectors, $Ax - b$ is an affine mapping of x and $\geq_{\mathcal{C}}$ is the the partial ordering induced by a pointed convex cone $\mathcal{C}$. To do so, we choose $Ax-b$ such that $(Ax-b)_1 = t$ and $(Ax-b)_{i+1} = (d_i - s(m_i))/\sigma_i$. Then, if $Ax - b$ is inside the quadratic cone Q defined as

$$Q := \left\{(q_1, q_2) \in \mathbb{R} \times \mathbb{C}^{d-1} | \, q_1 \geq \|q_2\|\right\},$$

we have

$$t = (Ax-b)_1 \geq \left\|\overrightarrow{(Ax-b)_{i+1}}\right\| = \sqrt{\sum_i \left(\frac{d_i - s(m_i)}{\sigma_i}\right)^2},$$

which implies equation (4.6). Thus (4.5) is equivalent to a minimization of a linear objective, constrained to the intersection of an affine plane $Ax - b$ with the cone Q.

In the case of incomplete tomography data, the quantum system s for which the minimum (4.5) is attained need not be unique, i. e., several quantum systems s may lead to the same value of t in (4.6). As one is usually interested in the performance of the laboratory system to engineer a particular quantum system, we ask for the best-case fidelity and worst-case fidelity with the designated system. Fortunately, the design is often a pure state or a unitary quantum gate. That is, it is an extreme point in the set of all states or the set of all channels, respectively. The fidelity with extreme points of a convex set can be written as linear objective. Thus we end up with a three step postprocessing:

1. Compute the optimal t in (4.6).
2. Compute the minimum fidelity consistent with that t. (4.8)
3. Compute the maximum fidelity consistent with that t.

Given a linear figure of merit, i. e., a linear fidelity, all steps are conic programs and thus reliably solvable numerically. This can be extended to quadratic fidelities, if the same quadratic cone trick applies. Also, robustness analysis of the conic constraints can give insights in how to improve the quality of the quantum system and the tomography.

4.2 States

We assume a finite dimensional Hilbert space $\mathcal{H}$. States are therefore given by density operators $\rho \in \mathcal{B}(\mathcal{H})$. They lie in the cone of semidefinite complex matrices S (i. e., $\rho \geq 0$). The normalization condition $\mathrm{tr}(\rho) = 1$ translates to two conic

constraints (i.e., $\operatorname{tr}(\rho) - 1 \geq 0$ and $-\operatorname{tr}(\rho) + 1 \geq 0$). The least squares fit (4.5) of a state ρ to given measurement results a_i with standard deviation σ_i for the operators A_i thus becomes a conic program in (t, ρ):

$$\begin{aligned} \text{minimize} \quad & t \\ \text{subject to} \quad & t^2 \geq \sum_i \left(\frac{a_i - \operatorname{tr}(\rho A_i)}{\sigma_i} \right)^2, \\ & \rho \geq 0, \\ & \operatorname{tr}(\rho) - 1 \geq 0, \\ & -\operatorname{tr}(\rho) + 1 \geq 0. \end{aligned} \tag{4.9}$$

Hence, it is a linear optimization problem in the cone $\mathcal{C} = Q \times S \times \mathbb{R}_+ \times \mathbb{R}_+$.

In the case that the designated state ψ is pure, the minimal and maximal fidelity with ψ can be calculated for fixed t via a conic program in ρ:

$$\begin{aligned} \text{minimize} \quad & \pm\langle\psi|\rho|\psi\rangle \\ \text{subject to} \quad & t^2 \geq \sum_i \left(\frac{a_i - \operatorname{tr}(\rho A_i)}{\sigma_i} \right)^2 \\ & \rho \geq 0, \\ & \operatorname{tr}(\rho) - 1 \geq 0, \\ & -\operatorname{tr}(\rho) + 1 \geq 0. \end{aligned} \tag{4.10}$$

4.3 Channels

Due to the duality between quantum channels and states on a larger Hilbert space [65, 64], the problem of finding a channel that fits best to given measurement results is essentially the same as for states.

A channel in the Heisenberg picture is a completely positive, unital map. As in Definition 3.2.3 on page 31 in the previous chapter, we associate an operator $\widetilde{T} \in \mathcal{B}(\mathcal{L}^2(\mathcal{H}_1, \mathcal{H}_2))$ to every channel $T\colon \mathcal{B}(\mathcal{H}_1) \to \mathcal{B}(\mathcal{H}_2)$ by

$$\widetilde{T}(x) := \sum_{k,l} T(|k\rangle\langle l|) x |l\rangle\langle k|.$$

Here $\mathcal{H}_1$ and $\mathcal{H}_2$ are finite dimensional Hilbert spaces of possibly different dimension and $|k\rangle$ denote basis vectors of $\mathcal{H}_1$. From Proposition 3.2.4 on page 32 we know that the channel T is completely positive if and only if $\widetilde{T}$ is positive semidefinite (i. e., $\widetilde{T} \in S$). Furthermore, Lemma 3.2.5 states that T is unital if and only if the partial trace $\operatorname{tr}_{\mathcal{H}_1} \widetilde{T} = \mathbb{1}$ (i. e., $\operatorname{tr}_{\mathcal{H}_1} \widetilde{T} - \mathbb{1} \in S$ and $-\operatorname{tr}_{\mathcal{H}_1} \widetilde{T} + \mathbb{1} \in S$). If $\widetilde{T}$ is known, the associate channel T is given by equation (3.13),

$$T(A) = \sum_{b,l} \widetilde{T}(|b\rangle\langle l|) A |l\rangle\langle b|. \tag{4.11}$$

The only difference for channels $T_*\colon \mathcal{B}(\mathcal{H}_1) \to \mathcal{B}(\mathcal{H}_2)$ is that the partial trace is over the output Hilbert space $\mathcal{H}_2$ instead of $\mathcal{H}_1$.

The least squares fit (4.5) of a channel T to given measurement results a_i with standard deviations σ_i for operators A_i given the input states ρ_i becomes:

$$\begin{aligned} &\text{minimize} && t \\ &\text{subject to} && t^2 \geq \sum_{i,b,l} \left(\frac{a_i - \operatorname{tr}(\rho_i \widetilde{T}(|b\rangle\langle l|) A_i |l\rangle\langle b|)}{\sigma_i} \right)^2, \\ & && \widetilde{T} \geq 0, \\ & && \operatorname{tr}_{\mathcal{H}_1}(\widetilde{T}) - \mathbb{1} \geq 0, \\ & && -\operatorname{tr}_{\mathcal{H}_1}(\widetilde{T}) + \mathbb{1} \geq 0. \end{aligned} \tag{4.12}$$

Thus, it is a linear optimization problem in $(t, \widetilde{T})$ with the cone $\mathcal{C} = Q \times S \times S \times S$.

Again, the fitted channel may not be unique. If the model system is a channel given by a unitary operator U, $\mathrm{ad_U(A)} := \mathrm{U^*AU}$, one can compute the minimum fidelity and maximum fidelity with the model for given t as above. The fidelity is taken as the channel fidelity of Definition 3.1.6 on page 27 for the concatenation of T with the inverse of the unitary channel,

$$F_U(T) := \langle \Omega | (\mathrm{id} \otimes UTU^*) | \Omega \rangle. \tag{4.13}$$

Here $|\Omega\rangle$ denotes the standard maximally entangled unit vector and id is the identity channel. This unitary channel fidelity is linear in T, and equal to 1 if and only if $T = \mathrm{ad_U}$. For unitary channels, the postprocessing is therefore done in three steps as for pure states.

The best-case fidelity and worst-case fidelity for the tomography of a unitary gate is particularly interesting as often the tomography is done under the assumption that the engineered system is indeed unitary. This results in an incomplete tomography if the system is considered as a general channel instead, as a general Stinespring isometry has far more parameters than a unitary operator on the input system. Typically, the assumption of having a unitary operation in the laboratory cannot be justified, as interactions with the environment cannot be precluded. In this case, the described scheme results in a fit of general and not necessarily unitary channels that are consistent with the measurement results, and the best-case fidelity and worst-case fidelity show how close channels that are consistent with the measurement results are to the designated unitary channel.

4.4 Implementation

The implementation was done using the Matlab package SeDuMi [47]. SeDuMi is a numerical solver for conic problems. The primal form is

$$\min \{\mathrm{Re}\,\langle c|x\rangle \,|Ax = b, x \in \mathcal{C}\}\,.$$

Here x is the vector to optimize, $Ax = b$ is an affine constraint and $\mathcal{C}$ is a symmetric cone, i. e., $\mathcal{C}$ is self-dual and homogeneous[1]. The dual form is

$$\max \{\mathrm{Re}\,\langle b|y\rangle \,|c - A^*y \in \mathcal{C}\}\,.$$

For matrices, the vectors x and y are obtained by column stacking, a technique also known as vectorization.

In this case, the formulation in the dual form was chosen and the code directly implements (4.9) and (4.12) and the corresponding fidelity optimizations. For the dual problem, SeDuMi's definition for the matrix cone is $S = \{x \,|x + x^* \geq 0\}$, so in order to ensure $x \geq 0$ we have to demand $x \in S$ and $x^* = x$, which splits into $ix \in S$ ($ix \geq ix^*$) and $-ix \in S$ ($ix^* \geq ix$). The Matlab code is given in Appendix A on page 151.

Note that writing equality constraints as two inequality constraints can lead to numerical difficulties, although this was not observed in the given implementation. An alternative approach is to use the primal form with explicit equality constraints and model the inequality constraints via so called slack variables s, that is, we use the fact that $Ax - b \geq 0$ if and only if there exists an $s \geq 0$ such that $Ax - s - b = 0$. The disadvantage of this approach is that we need a slack variable for every least squares term, and the number of terms is exponentially increasing in the number of qubits of the system in the case of complete tomography. Another alternative is to enforce the equality constraints via the choice of parameters to optimize, as, for example, done by Audenaert and De Moor [63]. This approach may destroy sparsity of A, and thus may require longer solution times.

For less sophisticated optimization packages than SeDuMi, the problem (4.5) can be formulated solely in terms of positive semidefinite cones, that is, without an explicit quadratic cone. In order to achieve this, we use the Shur complement [45], which states that for a hermitian matrix M partitioned as

$$M = \begin{pmatrix} A & B \\ B^* & C \end{pmatrix},$$

[1] Homogeneous means that for every non-zero element of the cone there exists a bijective map that maps this element to unity. This property about the scaling behavior of the cone is required for some solver algorithms. See [39] for a precise definition.

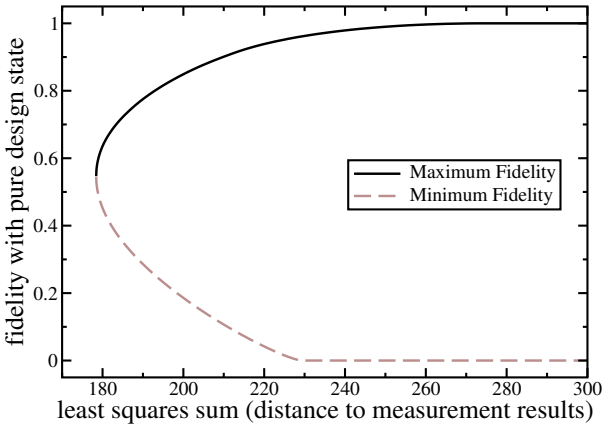

Figure 4.1: Result of least squares fit of imaginary tomography data for a four-qubit Dicke state with two excitations. The upper curve shows the best possible pure state fidelity of a fitted state with the Dicke state depending on the value of the least squares sum. The lower curve shows the worst-case fidelity of a fitted state. A fitting algorithm that does not guarantee to have found the global minimum of the least squares sum can result in any fidelity between these two curves.

where A and C are square, $M \geq 0$ is equivalent to $A \geq 0$ and $C - B^* A^{-1} B \geq 0$. For example, to write the problem (4.9) as a semidefinite program, we introduce for every pair (A_i, a_i) an auxiliary variable t_i and the affine mapping $(\rho, t_i) \mapsto M_i$ with

$$M_i = \begin{pmatrix} t_i & (\mathrm{tr}(\rho A_i) - a_i) \\ (\mathrm{tr}(\rho A_i) - a_i) & 1 \end{pmatrix} .$$

Looking at the Shur complement of M_i we see that M_i is positive semidefinite if and only if $t_i \geq (\mathrm{tr}(\rho A_i) - a_i)^2$. Minimizing the t_i with the constraints $M_i \geq 0$ and $\mathrm{tr}(\rho) = 1$ results in the desired least square fit density operator ρ. However, this requires an auxiliary variable t_i for every term of the least squares sum instead of a single auxiliary variable as in the formulation (4.5). Thus it leads to an exponential overhead in the number of qubits in the case of complete tomography.

4.5 Example

As example, we consider the four-qubit Dicke state with two excitations,

$$|\psi\rangle = \frac{1}{6} \left(|0011\rangle + |0101\rangle + |1001\rangle + |0110\rangle + |1010\rangle + |1100\rangle . \right)$$

We will use imaginary tomography data as it could have been obtained by an experiment such as [79]. That is, let ρ_{exp} be the engineered state, then the tomography data corresponds to the measurements

$$\mathrm{tr}(\rho_{\mathrm{exp}} (\sigma_i \otimes \sigma_j \otimes \sigma_k \otimes \sigma_l)) \tag{4.14}$$

for all 81 combinations of Pauli matrices, i. e., $i, j, k, l \in \{\mathrm{x}, \mathrm{y}, \mathrm{z}\}$. In the experiment [79], the statistics are gathered with single photon detectors connected to a multi-channel coincidence unit. Each setting (4.14) has $2^4 = 16$ possible four-fold coincidences. The rates of these coincidences are scaled according to the detection efficiencies, so the errors, e.g., due to the Poisson distribution, get scaled as well.

The result of the least squares fit is shown in Figure 4.1. It shows the maximum and minimum fidelity $\langle\psi|\rho_{\mathrm{fit}}|\psi\rangle$ depending on the value of the least squares sum (4.5),

$$\sum_{i,j,k,l=1}^{3} \left(\frac{d_{(i,j,k,l)} - \mathrm{tr}(\rho_{\mathrm{fit}}(\sigma_i \otimes \sigma_j \otimes \sigma_k \otimes \sigma_l))}{\sigma_i^2} \right)^2 ,$$

where we estimated the variances σ_i^2 by the measured number of clicks $\sigma_i^2 = n_i$, since the normal distribution approximation of a Poisson distribution has the same mean and variance as the Poisson distribution itself. However, for $n_i = 0$ this leads to a division by zero problem, so in this case, one has to choose a sufficiently large weight of the corresponding least squares term. Another choice of the variances would be $\sigma_i^2 = (N \,\mathrm{tr}(\rho_{\mathrm{fit}} A_i))$, where N is the total number of obtained coincidence clicks in the experiment containing the measurement operator A_i. However, a conic formulation (4.6) with the upper bound as linear objective would not be possible in this case.

The fidelity of 0.55 at the minimum of the least squares sum is of no particular interest for this imaginary experiment, however, what is interesting are the characteristics of the curves. Already a one percent larger least squares sum leads to a maximum fidelity of 0.64 and the minimum fidelity of 0.45, that is, the value of the fidelity changes by about 16 percent. For a 10 percent larger least squares sum the maximum fidelity is 0.82 and the minimum fidelity 0.22. A fitting algorithm that does not guarantee to have found the global minimum of the least squares sum[2] can in principle get stuck in a particular flat region of that sum and result in any fidelity between the two curves for the maximum and minimum fidelity.

4.6 Conclusion

Postprocessing tomography data via conic programming provides a certified fidelity for the state or the channel in the experiment. In the case of incomplete tomography of pure states or unitary channels, it calculates the minimum fidelity and maximum fidelity with respect to the designated quantum system, where the extrema are taken over all systems that are consistent with the observed data. The example shows that the minimum fidelity and maximum fidelity can change by large amounts, if the fit is allowed to have a least squares sum that is slightly larger than its global minimal

[2]For example, compare with the Powell algorithm proposed in [81].

value. In contrast to other methods, conic programming is guaranteed to find the global minimum of the least squares sum, and therefore neither overestimates nor underestimates the fidelity of the fit.

Chapter 5

Entanglement Estimation in Experiments

Entanglement is a fundamental resource for communication and computation that is unique to quantum mechanical systems. Therefore, the amount of entanglement inherent in a quantum state is of great experimental and theoretical interest. A standard method for establishing experimentally, whether a given state is entangled, is the measurement of an entanglement witness. An entanglement witness is an operator that is designed in such a way that a negative expectation value can only be obtained from an entangled state, while no conclusion about the given entanglement can be drawn from a positive expectation value. In contrast to a complete tomography combined with the computation of the value of an entanglement measure for the fitted state, this approach has the advantage that only few measurements are necessary to witness the entanglement. Often, only a single witness is measured for this purpose. However, the numerical value of the observed expectation value is usually considered to be of no further importance. In particular, no conclusions about the amount of entanglement are drawn, that is, witnesses are solely used for the detection of entanglement.

In this chapter, we show how a lower bound on a generic entanglement measure can be derived from the measured expectation values of any finite collection of entanglement witnesses. Thus, witness measurements provide quantitative information about the given entanglement without any additional experimental effort. More generally, the method uses the Legendre transform to characterize the best possible lower bound on any convex function $E(\rho)$ with compact sublevel sets for all states ρ that are consistent with given expectation values of a finite set of hermitian operators. The results in this chapter are prepublished in [4], together with an application to the multiphoton experiment [82]. Similar ideas and conclusions are reached by Eisert *et al.* [83].

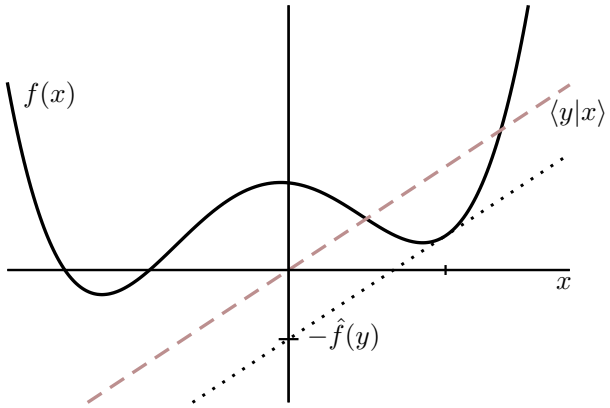

Figure 5.1: Legendre transform of the function $f(x)$ for a fixed value y. The Legendre transform $\hat{f}(y)$ is defined as the maximum gap between the line $\langle y|x\rangle$ (dashed line) and $f(x)$ (solid line). If f is differentiable, the maximum is attained at a value of x for which $f'(x) = y$, as shown by the affine function $\langle y|x\rangle - \hat{f}(y)$ (dotted line).

In [84], the Legendre transformation was used to characterize additivity properties of entanglement measures. The question how to estimate entanglement in the incomplete tomography setting was first addressed in [85]. Bounds on some entanglement measures from special Bell inequalities or entanglement witnesses have been obtained in [86, 87]. Methods to estimate entanglement measures in experiments by making measurements on several copies of a state have been discussed in [88, 89, 90].

5.1 Bound Construction

The bound construction arises from the theory of the Legendre transform [39, 41], which is also known as conjugate function, Fenchel transformation or dual function[1]. The Legendre transformation of a function f is defined to be the maximum gap between the line $\langle y|x\rangle$ and $f(x)$.

5.1.1 Definition (Legendre Transformation)**.** Let $f\colon \mathbb{R}^n \to \mathbb{R}$. Then the *Legendre transform* of f is defined as $\hat{f}\colon \mathbb{R}^n \to \mathbb{R}$,

$$\hat{f}(y) = \sup_{x \in \operatorname{dom} f} \{\langle y|x\rangle - f(x)\}\,, \tag{5.1}$$

where $\operatorname{dom} f$ denotes the domain of f. The domain of the Legendre transform consists of $y \in \mathbb{R}^n$ for which the supremum is finite.

Figure 5.1 shows the situation for a fixed y. As pointwise supremum of affine functions, the Legendre transform $\hat{f}$ is a convex function, no matter whether or not the

[1] The definitions may slightly differ. For example, [39] uses the term Legendre transform when they refer to differentiable functions and conjugate function for the general case.

function f is convex. Furthermore, if f is convex and closed[2], we have $\hat{\hat{f}} = f$ [39, 41]. The defining equation (5.1) leads to the inequality

$$f(x) + \hat{f}(y) \geq \langle y|x\rangle, \tag{5.2}$$

which is called Fenchel's inequality or Young's inequality.

The main idea is as follows. Let ρ be the given quantum state and $E(\rho)$ be the function to estimate. In the simplest case we have a single expectation value $w = \mathrm{tr}(\rho W)$ from the measurement of the hermitian operator W. So if we know the Legendre transform of E, Fenchel's inequality (5.2) already leads us to the lower bound

$$E(\rho) \geq \langle W|\rho\rangle - \hat{E}(W) = \mathrm{tr}(\rho W) - \hat{E}(W) = w - \hat{E}(W).$$

Observe that this is not the only bound we get from the expectation value w. Since every scalar multiple of the hermitian operator W is also a valid measurement, we have

$$E(\rho) \geq rw - \hat{E}(rW)$$

for every scalar r. So the optimal lower bound we obtain this way, let us call it $\tilde{\varepsilon}(w)$, is

$$E(\rho) \geq \tilde{\varepsilon}(w) = \sup_r \left\{ rw - \hat{E}(rW) \right\}, \tag{5.3}$$

which is itself a Legendre transformation of the function $r \mapsto \hat{E}(rW)$.

Note that the bound (5.3) does not require convexity of E. However, if we assume that E is indeed convex and has compact sublevel sets, that is, the sets $S_t^E = \{\rho \,|E(\rho) \leq t\}$ are compact, we obtain the best possible lower bound on E for the given measurement result. To see this, let us consider n hermitian operators $W_1, \ldots W_n$, so we also extend the analysis to the case of several measurements. Now let E be a convex function with compact sublevel sets that maps quantum states to real numbers. Furthermore, let $w_1, \ldots, w_n$ be the expectation values for the state ρ obtained in the experiment, i. e., $\mathbb{R} \ni w_i = \mathrm{tr}(\rho W_i)$, $i = 1, \ldots, n$. The best lower bound on $E(\rho)$ is then given by $\varepsilon\colon \mathbb{R}^n \to \mathbb{R}$,

$$\varepsilon(w_1, \ldots, w_n) = \inf_\sigma \{E(\sigma) \,|\mathrm{tr}(\sigma W_i) = w_i, i = 1, \ldots, n\}, \tag{5.4}$$

the infimum of E over all states that lead to the given expectation values.

5.1.2 Proposition. *If E is convex and has compact sublevel sets, then ε in* (5.4) *is convex and closed on*

$$\mathrm{dom}\,\varepsilon = \{\vec{w} \in \mathbb{R}^n \,|\exists \rho : \mathrm{tr}(\rho W_i) = w_i\}.$$

[2] A function f is closed if the epigraph $\{(x,t) \,|x \in \mathrm{dom}\, f,\ f(x) \leq t\}$ is closed, or equivalently, if the sublevel sets $\{x \in \mathrm{dom}\, f \,|f(x) \leq t\}$ are closed for all t.

Proof. The function (5.4) is convex on

$$\operatorname{dom}\varepsilon = \{\vec{w} \in \mathbb{R}^n \,|\exists\rho : \operatorname{tr}(\rho W_i) = w_i\}\,.$$

As the infimum always exists on $\operatorname{dom}\varepsilon$, we can find ρ_α for every $\vec{w}_\alpha \in \operatorname{dom}\varepsilon$ such that $\varepsilon(\vec{w}_\alpha) = E(\rho_\alpha) - \delta$ for arbitrary small δ. Let $\rho = \sum_\alpha \lambda_\alpha \rho_\alpha$ be a convex combination of such ρ_α. Then we have $\operatorname{tr}(\rho W_i) = \lambda_\alpha w_{\alpha_i}$, and due to convexity of E we get

$$E(\rho) \leq \sum_\alpha \lambda_\alpha E(\rho_\alpha) = \sum_\alpha \lambda_\alpha \varepsilon(\vec{w}_\alpha) + \delta.$$

With $\delta \to 0$ and since $\varepsilon(\sum_\alpha \lambda_\alpha \vec{w}_\alpha) \leq E(\rho)$ we conclude that ε is convex.

Furthermore, ε in (5.4) is closed. The sublevel sets of ε are given by

$$S_t^\varepsilon = \{\vec{w} \,|\forall\delta > 0 \exists\rho : \operatorname{tr}(\rho W_i = w_i), i = 1, \ldots, n; E(\rho) \leq t + \delta\}\,,$$

where we inserted the definition of the infimum. This can be rewritten using the sublevel sets S_t^E of E, that we assume to be compact, leading to

$$\begin{aligned} S_t^\varepsilon &= \bigcap_{\delta>0} \left\{\vec{w} \,\middle|\exists\rho : \operatorname{tr}(\rho W_i = w_i), i = 1, \ldots, n; \rho \in S_{t+\delta}^E\right\} \\ &= \bigcap_{\delta>0} \left\{\operatorname{tr}(\rho\vec{W}) \,\middle|\rho \in S_{t+\delta}^E\right\}. \end{aligned}$$

Here $\operatorname{tr}(\rho\vec{W})$ is the vector with i-th component $\operatorname{tr}(\rho W_i)$. Since $\operatorname{tr}(\rho\vec{W})$ is continuous and $S_{t+\delta}^E$ compact, S_t^ε is an intersection of compact sets and therefore compact. In particular, we have that ε is closed. ∎

Because ε in (5.4) is convex and closed, we have $\hat{\hat{\varepsilon}} = \varepsilon$. Therefore we can characterize $\varepsilon\colon \mathbb{R}^n \to \mathbb{R}$ as the supremum of all affine functions below it. So consider bounds of the type

$$\varepsilon(\vec{w}) \geq \langle\vec{r}|\vec{w}\rangle - c \tag{5.5}$$

for arbitrary $\vec{r} \in \mathbb{R}^n$ and $c \in \mathbb{R}$, and $\langle\vec{r}|\vec{w}\rangle = \sum_i r_i w_i$. Note that by definition of ε in equation (5.4) this is the same as saying that $E(\rho) \geq \langle\vec{r}|\vec{w}\rangle - c$ for every ρ giving the expectation values w_i, $i = 1, \ldots, n$. As a constant, c does not depend on $\vec{w}$, so

$$c \geq \langle\vec{r}|\vec{w}\rangle - E(\rho)$$

has to hold for any $\vec{w}$ for which a state ρ exists such that $\vec{w} = \operatorname{tr}(\rho\vec{W})$, and therefore for any ρ. The best choice for c is thus

$$\begin{aligned} c &= \sup_\rho \left\{\sum_i r_i \operatorname{tr}(\rho W_i) - E(\rho)\right\} \\ &= \sup_\rho \left\{\langle\rho| \sum_i r_i W_i\rangle - E(\rho)\right\} = \hat{E}\left(\sum_i r_i W_i\right), \end{aligned} \tag{5.6}$$

where we used the condition $\vec{w} = \mathrm{tr}(\rho\vec{W})$ and the linearity of the trace. So for a given slope $\vec{r}$, the optimal constant c only depends on the operator $\mathcal{W} = \sum_i r_i W_i$ and the Legendre transform of E,

$$\hat{E}(\mathcal{W}) = \sup_{\rho} \left\{ \mathrm{tr}(\rho\mathcal{W}) - E(\rho) \right\}. \tag{5.7}$$

We can now use the optimal constant c in (5.6) for the formula (5.5) to write $\varepsilon(\vec{w})$ as the supremum of all affine functions below it,

$$\varepsilon(\vec{w}) = \sup_{(r_1,\ldots,r_n)} \left\{ \sum_j r_j w_j - \hat{E}\left(\sum_i r_i W_i \right) \right\}, \tag{5.8}$$

that is, as Legendre transform of $\hat{\varepsilon}(\vec{r}) = \hat{E}(\sum_i r_i W_i)$. So in this case, the Legendre transform (5.8) is indeed the optimal lower bound by definition of $\varepsilon(\vec{w})$ in (5.4).

In summary, we obtained the following result.

5.1.3 Theorem. *Let $w_1, \ldots, w_n$ be the expectation values of the hermitian operators $W_1, \ldots, W_n$ for a given quantum state ρ, i. e., $w_i = \mathrm{tr}(\rho W_i)$, $i = 1, \ldots, n$. Furthermore, let $E\colon \mathcal{B}(\mathcal{H}) \to \mathbb{R}$ be a function that maps quantum states to real numbers. Then the best lower bound on $E(\rho)$,*

$$\varepsilon(w_1, \ldots, w_n) = \inf_{\sigma} \left\{ E(\sigma) \,|\mathrm{tr}(\sigma W_i) = w_i, i = 1, \ldots, n \right\}$$

is bounded below by

$$\tilde{\varepsilon}(w_1, \ldots, w_n) = \sup_{(r_1,\ldots,r_n)} \left\{ \sum_j r_j w_j - \hat{E}\left(\sum_i r_i W_i \right) \right\}.$$

Moreover, if E is convex and E has compact sublevel sets, then $\varepsilon = \tilde{\varepsilon}$.

The computation of $\tilde{\varepsilon}(w_1, \ldots, w_n)$ does involve the computation of the two Legendre transforms, $\hat{\tilde{\varepsilon}}$ (equation (5.8)) and $\hat{E}(\mathcal{W})$ (equation (5.7)). For a fixed $\vec{r}$, the constant $c = \hat{E}\left(\sum_i r_i W_i\right)$ already gives the best lower bound of the form (5.5). So even if the optimization over $\vec{r}$ does not attain the true global optimal value, for example, due to numerical difficulties, the result is still a lower bound on E. This situation is depicted in Figure 5.2. By contrast, we see from equation (5.8) that an error in the computation of $\hat{E}$ can lead to a lower bound that is too optimistic.

Typically n is small, for example, $n = 1$, compared to the dimension of the space of all Hermitian operators on the given Hilbert space. So (5.8) is an optimization over a low-dimensional space. The ability to efficiently compute $\hat{E}(\sum_i r_i W_i)$ depends on the entanglement measure E and witnesses W_i chosen. Below, this is discussed for the entanglement of formation and the geometric measure of entanglement.

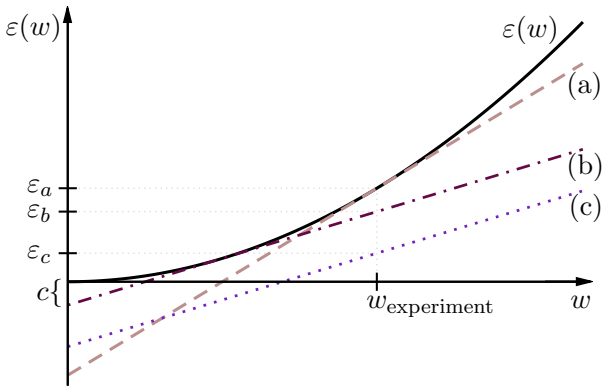

Figure 5.2: A schematic view of the lower bound construction for the convex function $E(\rho)$ in case of a single measured expectation value $w_{\text{experiment}}$. The figure shows $\varepsilon(w)$, the optimal lower bound on E over all states conforming to the expectation value as function of w, and lower bounds given by affine functions $rw - c$ below it, as in (5.5). The line (a) corresponds to the optimal solution $\varepsilon_a = \varepsilon(w_{\text{experiment}})$ for the slope r and constant c for $w_{\text{experiment}}$, that is, r is the slope for which the supremum (5.8) is attained at the point $w_{\text{experiment}}$. The line (b) shows an affine function with optimal c but suboptimal r leading to the lower bound ε_b, whereas the line (c) has neither optimal slope r, nor optimal constant c, leading to the bound ε_c.

5.2 Entanglement Measures

One of the main resources in quantum information is entanglement, which is used in quantum algorithms such as teleportation or dense coding[3]. A bipartite state is entangled if its correlation cannot be explained with a classical random generator, that is, if it is not separable [91, 32]. Deciding whether or not a given state is entangled is one of the basic tasks of quantum information theory. In principle, this information can be deduced from the explicit form of the density operator. However, this procedure requires a complete tomography of the state, and hence the measurement of a number of matrix elements, which grows exponentially in the number of qubits of the system. It is therefore highly desirable to verify entanglement on the basis of only a few, maybe only one measurement.

Entanglement witnesses [24] are observables designed for this purpose. By definition, they have positive expectation on every separable state, so when a negative expectation is found in some state, it must be entangled. Consequently, entanglement witnesses have been used in many experiments [92, 93, 94, 95, 96, 82], and their theory is far developed [97, 98, 99, 86, 100]. Moreover, by the elementary duality theory of convex sets, for every entangled state there is a witness which has negative expectation on it. So we can indeed detect the entanglement of a state with a single

[3]See [35] for a description of these and other quantum algorithms.

measurement.

Besides the mere detection, the quantification of entanglement is an even more challenging problem in the field. Many entanglement measures have been introduced for this purpose (see [25]). Typically, their computation involves nontrivial optimizations over large sets, even if the density matrix of a state is fully known. Needless to say that their determination in experiments is a tremendous task.

With the above method, entanglement witnesses can not only be used for the detection of entanglement, but also for its quantification. Any measured negative expectation value of a witness turns into a nontrivial lower bound on entanglement measures. Below, the the procedures for computing such bounds are described in detail for the entanglement of formation [58] and the geometric measure of entanglement [101].

Note that the method also covers the case of incomplete tomography. For any finite set of measured expectation values, it characterizes the best possible bound on any convex entanglement measure consistent with these expectations.

5.2.1 Convex Roof Constructions

Many entanglement measures are defined by an extension process called convex roof construction [58], which extends a function E defined for pure states, $|\psi\rangle \mapsto E(|\psi\rangle)$, to a function for mixed states via

$$E(\rho) = \inf_{p_i, |\psi_i\rangle} \left\{ \sum_i p_i E(|\psi_i\rangle) \,\middle|\, 0 \leq p_i \leq 1, \sum_i p_i = 1, \sum_i p_i |\psi_i\rangle\langle\psi_i| = \rho \right\},$$

i. e., as infimum over all decompositions of ρ into extreme points of the convex state space. The convex roof construction is the minimal convex extension of the function. It is the largest convex underestimator of E on pure states, and the epigraph[4] of the convex roof is the convex hull of the epigraph of E. If the entanglement measure is defined via a convex roof construction, the calculation of the Legendre transform $\hat{E}$, equation (5.7), can be simplified to a variational problem over pure states only,

$$\begin{aligned} \hat{E}(\mathcal{W}) &= \sup_{\rho} \{\mathrm{tr}(\rho\mathcal{W}) - E(\rho)\} \\ &= \sup_{\rho} \left\{ \mathrm{tr}(\rho\mathcal{W}) - \inf_{p_i, |\psi_i\rangle} \sum_i p_i E(|\psi_i\rangle) \right\} \\ &= \sup_{p_i, |\psi_i\rangle} \left\{ \sum_i p_i \Big(\langle\psi_i|\mathcal{W}|\psi_i\rangle - E(|\psi_i\rangle) \Big) \right\} \\ &= \sup_{|\psi\rangle} \{\langle\psi|\mathcal{W}|\psi\rangle - E(|\psi\rangle)\}\,. \end{aligned} \tag{5.9}$$

[4] The epigraph of a function f, $f(x) \in \mathbb{R}$, is the set above the graph of f, $\{(x,t)\,|x \in \mathrm{dom}\, f,\, f(x) \leq t\,\}$.

Here we used the constraint $\rho = \sum_i p_i |\psi_i\rangle\langle\psi_i|$ and the fact that the optimal value is attained on an extreme point of the convex set of states, since the convex combinations of terms in the sum cannot be larger than the largest term.

Note that the amount of entanglement does not change under local unitary operations. Consequently, if we look at vectors of the form $|\psi\rangle = (U_1 \otimes U_2)|\phi\rangle$, with unitary operators U_1 and U_2 for fixed $|\phi\rangle$, then the optimization in (5.9) only involves the maximization of the first term $\langle\psi|\mathcal{W}|\psi\rangle$. For example, consider a witness operator of the form $\mathcal{W} = \alpha\mathbb{1} - |\chi\rangle\langle\chi|$. In this case, we have to maximize

$$|\langle\chi|\psi\rangle|^2 = |\langle\chi|(U_1 \otimes U_2)\phi\rangle|^2.$$

The maximum is attained when both vectors $|\chi\rangle$ and $|\psi\rangle$ have the same Schmidt basis, and the Schmidt coefficients are ordered in the same way [100]. Therefore, the Legendre transform (5.9) reduces to an optimization over Schmidt coefficients for the given Schmidt basis of $|\chi\rangle$. That is, we can reduce the set of states the supremum is taken over. For example, if $|\chi\rangle \in \mathcal{H} \otimes \mathcal{H}$, $\dim \mathcal{H} = d$, we only need to optimize d positive numbers instead of d^2 complex coefficients that would be required for the full state space.

5.2.2 Entanglement of Formation

Given the Hilbert space $\mathcal{H} \otimes \mathcal{K}$, the entanglement of formation E_{F} [58] is defined as the convex roof of

$$E_{\mathrm{vN}}(|\psi\rangle) = S(\mathrm{tr}_{\mathcal{H}}(|\psi\rangle\langle\psi|)) = S(\mathrm{tr}_{\mathcal{K}}(|\psi\rangle\langle\psi|)),$$

where $S(\rho) = -\mathrm{tr}(\rho \log_2 \rho)$ is the von Neumann entropy and $\mathrm{tr}_{\mathcal{H}}$ and $\mathrm{tr}_{\mathcal{K}}$ are the partial traces over the Hilbert spaces $\mathcal{H}$ and $\mathcal{K}$, respectively. The entanglement of formation can be interpreted as the least expected entanglement of any ensemble of pure states realizing ρ, or equivalently, the minimal entanglement that must be invested to realize the state in an experiment.

We will utilize the Gibb's variational principle (see [102]) from statistical mechanics for the computation of the variational problem (5.9) for the Legendre transform $\hat{E}_{\mathrm{F}}$. That is, we rewrite the entropy in terms of the Legendre transform of the free energy F,

$$S(\rho) = \inf_H \{\mathrm{tr}(\rho H) - F(H)\} = -\mathrm{tr}(\rho \log \rho). \tag{5.10}$$

Here the infimum is taken over all hermitian operators H. In comparison to the expression in statistical mechanics, we have set the Boltzmann constant and the inverse temperature to one. However, we will use the natural logarithm as in statistical mechanics, so one has to accordingly scale the obtained bound at the end. The free energy F is given by

$$F(H) = \inf_\rho \{\mathrm{tr}(\rho H) - S(\rho)\}, \tag{5.11}$$

where the infimum is taken over all states ρ. From statistical mechanics we known that this infimum is attained at

$$\rho = \frac{1}{\operatorname{tr}(e^{-H})} e^{-H}.$$

Inserting the optimal ρ into (5.11) leads us to

$$\begin{aligned} F(H) &= \operatorname{tr}\left(\frac{1}{\operatorname{tr}(e^{-H})} e^{-H} H\right) - S\left(\frac{1}{\operatorname{tr}(e^{-H})} e^{-H}\right) \\ &= \operatorname{tr}\left(\frac{1}{\operatorname{tr}(e^{-H})} e^{-H} H\right) + \operatorname{tr}\left(\frac{1}{\operatorname{tr}(e^{-H})} e^{-H} \log \frac{1}{\operatorname{tr}(e^{-H})} e^{-H}\right) \\ &= -\log\left(\operatorname{tr}\left(e^{-H}\right)\right). \end{aligned} \tag{5.12}$$

With this result, equation (5.10) can be written as

$$S(\rho) = -\operatorname{tr}(\rho \log \rho) = \inf_H \{\operatorname{tr}(\rho H) + \log(\operatorname{tr} e^{-H})\},$$

so we conclude that the infimum is attained for $H = -\log \rho$.

Inserting the Legendre transform expression of the entropy (5.10) into equation (5.9) for the Legendre transform $\hat{E}_{\mathrm{F}}$ we get

$$\begin{aligned} \hat{E}_{\mathrm{F}}(\mathcal{W}) &= \sup_{|\psi\rangle} \left\{ \langle\psi|\mathcal{W}|\psi\rangle - \inf_H \{\operatorname{tr}(\operatorname{tr}_{\mathcal{K}}(|\psi\rangle\langle\psi|)H) - F(H)\} \right\} \\ &= \sup_{|\psi\rangle} \sup_H \{\langle\psi|\mathcal{W}|\psi\rangle - \operatorname{tr}(\operatorname{tr}_{\mathcal{K}}(|\psi\rangle\langle\psi|)H) + F(H)\} \\ &= \sup_{|\psi\rangle} \sup_H \{\langle\psi|(\mathcal{W} - (H \otimes \mathbb{1}))|\psi\rangle + F(H)\}. \end{aligned} \tag{5.13}$$

Thus, $\hat{E}_{\mathrm{F}}$ is a joint supremum over pure states $|\psi\rangle \in \mathcal{H} \otimes \mathcal{K}$ and hermitian operators H on $\mathcal{B}(\mathcal{H})$.

The main virtue of rewriting $\hat{E}_{\mathrm{F}}$ in this form is that when one of the variables of the suprema is fixed, the supremum over the other variable can be calculated directly without the use of a search algorithm. For fixed H, the optimal $|\psi\rangle$ in (5.13) is an eigenvector for the largest eigenvalue of $(\mathcal{W} - (H \otimes \mathbb{1})$. For fixed $|\psi\rangle$, the optimization (5.13) is equivalent to

$$\sup_H \{-\operatorname{tr}((\operatorname{tr}_{\mathcal{K}} |\psi\rangle\langle\psi|)H) + F(H)\} = -\inf_H \{\operatorname{tr}((\operatorname{tr}_{\mathcal{K}} |\psi\rangle\langle\psi|)H) - F(H)\}, \tag{5.14}$$

since the first term $\langle\psi|\mathcal{W}|\psi\rangle$ in (5.13) is constant. So the optimal H of (5.14) is the same as the optimal H for the Legendre transformation (5.10) with $\rho = \operatorname{tr}_{\mathcal{K}} |\psi\rangle\langle\psi|$, that is,

$$H = -\log(\operatorname{tr}_{\mathcal{K}} |\psi\rangle\langle\psi|).$$

And since the suprema in (5.13) commute, by alternating between them we monotonically increase the obtained value for $\hat{E}_{\mathrm{F}}(\mathcal{W})$ in every step and get a convergence to a local maximum.

If the local maximum is indeed a global one, the value of $\hat{E}_{\mathrm{F}}(\mathcal{W})$ can be used in (5.8) to obtain a lower bound on the entanglement of formation for the given measured expectation values. However, as the algorithm itself does not guarantee to have found a global optimum, the corresponding $\varepsilon(\vec{w})$ in (5.8) could overestimate the entanglement of formation for the given experimental state. On the other hand, in the example given in [4], the optimal value found was independent of the starting point, giving strong support to the claim of having found the global optimum.

5.2.3 Geometric Measure of Entanglement

The geometric measure of entanglement E_{G} is an entanglement monotone for multipartite systems [103, 104, 101]. For an n-partite system it is defined as the convex roof of

$$E_{\mathrm{G}}(|\psi\rangle) = 1 - \sup_{|\phi\rangle=|\varphi_1\rangle\otimes\cdots\otimes|\varphi_n\rangle} |\langle\phi|\psi\rangle|^2, \tag{5.15}$$

that is, as one minus the maximal squared overlap with a fully separable state. Thus, it can be interpreted as distance or angle to the nearest unentangled state. For pure states, the geometric measure is a lower bound on the relative entropy and one can derive from it an upper bound on the number of states which can be discriminated perfectly by local operations and classical communication [105, 106].

Inserting the definition (5.15) into (5.9), the Legendre transform $\hat{E}_{\mathrm{G}}$ becomes

$$\hat{E}_{\mathrm{G}}(\mathcal{W}) = \sup_{|\psi\rangle} \sup_{|\phi\rangle=|\varphi_1\rangle\otimes\cdots\otimes|\varphi_n\rangle} \{\langle\psi|(\mathcal{W} + |\phi\rangle\langle\phi|)|\psi\rangle - 1\}\,. \tag{5.16}$$

As in the case of entanglement of formation, we have a supremum over two variables, which we solve by alternately fixing one variable and optimizing the other. For a fixed $|\phi\rangle$, the supremum over $|\psi\rangle$ is attained if $|\psi\rangle$ is an eigenvector for the largest eigenvalue of the operator $(W + |\phi\rangle\langle\phi|)$. For a fixed vector $|\psi\rangle$, the term $\langle\psi|\mathcal{W}|\psi\rangle$ in (5.16) is fixed and we have to solve

$$\sup_{|\phi\rangle=|\varphi_1\rangle\otimes\cdots\otimes|\varphi_n\rangle} |\langle\psi|\phi\rangle|^2, \tag{5.17}$$

where the supremum is taken over all separable states $|\phi\rangle$. We do this optimization for every tensor factor $|\varphi_i\rangle$ separately. Suppose we want to optimize the i-th tensor factor $|\varphi_i\rangle$ while fixing the other tensor factors $|\varphi_{j\neq i}\rangle$. Then, we look at the Schmidt decomposition of $|\psi\rangle$ with respect to the partition into the Hilbert spaces $\mathcal{H}_i$ and $\mathcal{H}_{\{1,\ldots,n\}\setminus i}$. So let

$$|\psi\rangle = \sum_{\alpha} \lambda_\alpha |\alpha_i\rangle \otimes |\alpha_{\{1,\ldots,n\}\setminus i}\rangle$$

be the corresponding Schmidt decomposition after a proper reordering of tensor factors. Then we have

$$|\langle\psi|\phi\rangle|^2 = |\sum_{\alpha} \lambda_\alpha \langle\alpha_i|\varphi_i\rangle\langle\alpha_{\{1,\ldots,n\}\setminus i}|\varphi_{\{1,\ldots,n\}\setminus i}\rangle|^2,$$

where

$$|\varphi_{\{1,\ldots,n\}\setminus i}\rangle = \bigotimes_{j\in\{1,\ldots,n\}\setminus i} |\varphi_j\rangle.$$

This is maximal, if the vectors are parallel. Therefore, the supremum (5.17) for fixed $|\varphi_j\rangle$, $j \in \{1,\ldots,n\}\setminus i$, is attained at

$$|\varphi_i\rangle = N \sum_\alpha \lambda_\alpha \langle \alpha_{\{1,\ldots,n\}\setminus i} | \varphi_{\{1,\ldots,n\}\setminus i}\rangle |\alpha_i\rangle,$$

where N denotes a normalization such that $\|\varphi_i\| = 1$. We do this optimization for all $|\varphi_i\rangle$, $i = 1,\ldots,n$, successively

Since the suprema commute, by alternating the optimizations for $|\psi\rangle$ and $|\varphi_i\rangle$, $i = 1,\ldots,n$, we are monotonically increasing the obtained value for $\hat{E}_{\mathrm{G}}(\mathcal{W})$ arriving at a local maximum. However, the same special note of warning as for the algorithm for the entanglement of formation applies. If the final value for $(|\psi\rangle, |\phi\rangle)$ does not belong to a global optimum, the optimization (5.8) suggests a lower bound on E_{G} that is too large. On the other hand, for witness operators of the form $\alpha\mathbb{1} - |\chi\rangle\langle\chi|$, the Legendre transform $\hat{E}_{\mathrm{G}}$ can be calculated analytically [4], so the lower bound property of (5.8) is ensured.

5.3 Conclusion

We developed a method that, based on given expectation values, provides a lower bound for a functional on states, if the Legendre transform of that functional is known. If the functional is convex and has compact sublevel sets, then the lower bound is optimal in the sense that it is the minimal value of the functional for all states that are compatible with the given expectation values. The method allows to estimate the amount of entanglement of a state using standard witness measurements, without requiring any additional experimental effort. This is interesting, as it means that the method can also be applied to past experiments, where witness measurements were merely used to detect entanglement qualitatively. The method is also well suited to provide bounds in the case of incomplete tomography. For the entanglement of formation and the geometric measure of entanglement, we have given numerical algorithms to compute the Legendre transformation.

Chapter 6

The Meaner King

A ship-wrecked physicist gets stranded on a far-away island that is ruled by a mean king who loves cats and hates physicists since the day when he first heard what happened to Schrödinger's cat. A similar fate is awaiting the stranded physicist. Yet, mean as he is, the king enjoys defeating physicists on their own turf, and therefore he maliciously offers an apparently virtual chance of rescue. He takes the physicist to the royal laboratory, a splendid place where experiments of any kind can be performed perfectly. There the king invites the physicist to prepare a certain silver atom in any state she likes. The king's men will then measure one of the three Cartesian spin components of this atom — they'll either measure σ_x, σ_y, or σ_z without, however, telling the physicist which one of these measurements is actually done. Then it is again the physicist's turn, and she can perform any experiment of her choosing. Only after she's finished with it, the king will tell her which spin component had been measured by his men. To save her neck, the physicist must then state correctly the measurement result that the king's men had obtained. Much to the king's frustration, the physicist rises to the challenge — and not just by sheer luck: She gets the right answer any time the whole procedure is repeated. How does she do it?

— *The King's Problem* [26]

The above tale, known as the *mean king problem*, describes the basic quantum mechanical retrodiction problem to use an entangled copy of a quantum system to reconstruct the values of measurements, which can no longer be obtained from the system itself (see Figure 6.1). In the above form, the problem was solved by Vaidman, Aharonov, and Albert [30]. The solution relies on the fact that the three Pauli bases the king's men are allowed to measure have the special property of being mutually unbiased.

6.0.1 Definition (Mutually Unbiased Bases)**.** A number of k orthonormal bases $\{\Phi_a(i)\}$, $a = 1, \ldots, k$, on a given Hilbert space $\mathcal{H}$ with dimension d are mutually unbiased, if

$$|\langle\Phi_a(i)|\Phi_b(j)\rangle|^2 = \frac{1}{d}$$

for all $a \neq b$. Here, $\Phi_a(i)$ is the i-th vector of the a-th basis.

Mutually unbiasedness ensures that Alice's measurement result is totally useless for her final guess, if Alice is making her measurement in one of the king's bases, but didn't choose the same basis as the king's men. This even holds if the game is played many times and she is allowed to collect statistical data. Therefore, mutually unbiased bases seem to be a particular mean choice for the king and the generalization of the problem is usually done in the following way (see Ref. [28] for variations of the problem):

6.0.2 Problem (Mean King)**.** Alice is supposed to retrodict the outcome of a von Neumann measurement of a basis that is randomly chosen by the king's men from a set of known mutually unbiased orthonormal bases $\{\Phi_a(i)\}$. Here $\Phi_a(i) \in \mathcal{H}$ is the i-th vector of the a-th basis, $i = 1, \ldots, d = \dim \mathcal{H}$, $a = 1, \ldots, k$. To do so, Alice prepares a quantum state $\rho \in \mathcal{B}(\mathcal{H} \otimes \mathcal{K})$. She sends the first part of the state to the king's men. The king's men randomly choose a basis a and make a von Neumann measurement in that basis, but keep their result i secret. The quantum system is returned to Alice and she is allowed to do a final measurement $\{F_x\}$ on it. After that, the quantum system is discarded. She is told the choice a of the king's men. Based on her result x and the basis a, Alice guesses the result that the king's men obtained with her estimation function $s(a, x)$. She dies a cruel death, if $s(a, x) \neq i$. (See Figure 6.1.)

Let k be the number of bases of $\mathcal{H}$ with $\dim \mathcal{H} = d$. Then mutually unbiased bases are known to exist for $k \leq d + 1$ if d is the power of a prime [107]. For other dimensions, the existence of $k = d + 1$ mutually unbiased bases is still an open problem [108]. Nevertheless, if the king's men's choice is between mutually unbiased bases, there is always a solution to the retrodiction problem, where Alice obtains the result of the king's men's measurement with certainty [26, 109, 110, 27, 29]. An experimental realization was done by Schulz *et al.* [111].

6.1 The Meaner King Problem

As there always exists a solution for Alice, choosing between unbiased bases is not unkind at all. On the contrary, as we demand from Alice's solution that she is right in every single run, statistical correlations between the bases do not affect the

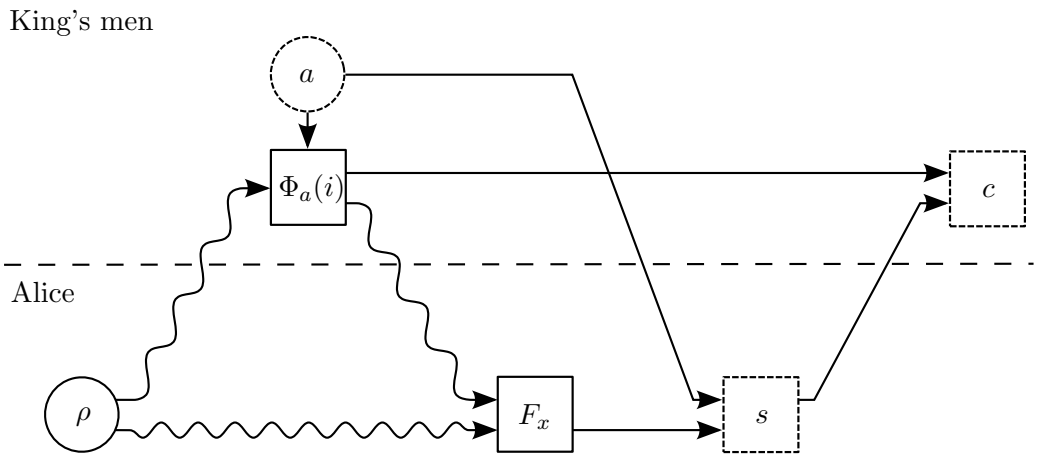

Figure 6.1: The mean king problem protocol: Alice prepares a quantum state ρ; The king's men choose a basis a among several choices and make a von Neumann measurement in that basis $\{\Phi_a(i)\}$; Alice is allowed to do a final measurement F_x afterwards; After that, she is told the king's men basis choice a; She has to guess the measurement result of the king's men (function s); Alice guess x and the king's measurement result i are compared (function c) and Alice dies a cruel death if they don't match (not depicted).

difficulty of the problem. We therefore drop the assumption of unbiasedness of the bases: The king chooses between any choice of finitely many bases.

6.1.1 Problem (Meaner King)**.** Alice is supposed to retrodict the outcome of a von Neumann measurement of a basis of the Hilbert space $\mathcal{H}$, $\dim \mathcal{H} = d$, that is randomly chosen by the king's men from a known and finite set of k bases. Otherwise, the procedure is the same as in problem 6.0.2 (see Figure 6.1).

Not very much has been done in the case without the assumption of mutually unbiasedness. Some special cases have been discussed in [112, 113, 29, 114].

With his choice $\{|\Phi_a(i)\rangle\}$ for the bases, the king determines how many operators he is able to distinguish with his measurements. This quantity is given by the dimension of the space

$$\mathcal{R} = \mathrm{span}_{\mathbb{R}} \left\{ |\Phi_a(i)\rangle\langle\Phi_a(i)| \,|i = 1, \dots d, a = 1, \dots, k \right\}, \tag{6.1}$$

the space of hermitian operators spanned by the projectors given by the bases.

6.1.2 Lemma. *The space $\mathcal{R}$ is at most $(k(d-1)+1)$-dimensional.*

Proof. Since $\sum_i |\Phi_a(i)\rangle\langle\Phi_a(i)| = \mathbb{1}$, we have one dimension shared by all bases and

therefore at most $(d-1)$ new dimensions for every additional basis. In summary this gives at most $(k(d-1)+1)$ dimensions for the space $\mathcal{R}$. ■

We will identify two important situations: Firstly, the case where there are no degeneracies and the dimension of $\mathcal{R}$ is indeed $(k(d-1)+1)$, i. e., maximal in the sense of Lemma 6.1.2. Secondly, the case where $\mathcal{R}$ is the space of all hermitian operators on the given Hilbert space, i. e., the dimension is maximal, $\dim \mathcal{R} = d^2$.

6.1.3 Definition (Non-Degenerate)**.** A set of bases $\{|\Phi_a(i)\rangle\}$ is *non-degenerate*, if the space spanned by the operators $|\Phi_a(i)\rangle\langle\Phi_a(i)|$ is $(k(d-1)+1)$-dimensional.

6.1.4 Definition (Tomographically Complete)**.** A set of bases is *tomographically complete*, if the operators $|\Phi_a(i)\rangle\langle\Phi_a(i)|$ span the whole space of hermitian operators on $\mathcal{H}$. Here $|\Phi_a(i)\rangle \in \mathcal{H}$ is the i-th vector of the a-th basis.

Restricting the king to non-degenerate bases includes the case of mutually unbiased bases but prevents degeneracies and complex dependencies, as we will see later.

6.1.5 Lemma. *Mutually unbiased bases are non-degenerate.*

Proof. Let $\{|\Phi_a(i)\rangle\}$, $i = 1, \dots d$, $a = 1, \dots, k$ be a set of k mutually unbiased bases, i. e.,

$$|\langle\Phi_a(i)|\Phi_b(j)\rangle|^2 = (1-\delta_{ab})\frac{1}{d} + \delta_{ab}\delta_{ij}. \tag{6.2}$$

The dimension of the real linear span of the projectors $|\Phi_a(i)\rangle\langle\Phi_a(i)|$ is given by the rank of the matrix

$$M_{ai,bj} := \langle|\Phi_a(i)\rangle\langle\Phi_a(i)|||\Phi_b(j)\rangle\langle\Phi_b(j)|\rangle_{\mathrm{HS}} = |\langle\Phi_a(i)|\Phi_b(j)\rangle|^2 = (1-\delta_{ab})\frac{1}{d} + \delta_{ab}\delta_{ij},$$

where $\langle A|B\rangle_{\mathrm{HS}} = \mathrm{tr}(A^*B)$ is the Hilbert Schmidt scalar product. The last equality is due to the condition (6.2). With the ordering

$$(a,i) = (1,1), \dots, (1,d), (2,1), \dots,$$

we have that δ_{ab} corresponds to an operator $\mathbb{1}_k \otimes E_d$, where E_d is the $d \times d$-matrix with all matrix elements equal to 1. From this it follows that

$$M = \frac{1}{d}(E_{kd} - \mathbb{1} \otimes E_d) + \mathbb{1} = \frac{1}{d}(E_k - \mathbb{1}) \otimes E_d + \mathbb{1}.$$

As the matrix $(E_k - \mathbb{1})$ has the simple eigenvalue $(k-1)$ and the $(k-1)$-fold eigenvalue -1, the matrix M has the eigenvalues:

$$\begin{aligned} 1 + (k-1) \qquad & \text{simple}, \\ 1 + -1 = 0 \qquad & (k-1)\text{-fold}, \\ 1 \qquad & k(d-1)\text{-fold}. \end{aligned}$$

Thus, the rank of M or equivalently the dimension of the real linear span of projectors $|\Phi_a(i)\rangle\langle\Phi_a(i)|$ is $k(d-1)+1$. In particular $k = d+1$ mutually unbiased bases are therefore tomographically complete. ■

The following lemma reveals why the case $k = d+1$ is of particular interest.

6.1.6 Lemma. *Let $\mathcal{H}$ be a Hilbert space, $\dim \mathcal{H} = d$, and S be a set of tomographically complete bases, $|S| = k$, with vectors $|\Phi_a(i)\rangle$, $i = 1, \ldots, d$, $a = 1, \ldots, k$. Then we have $k \geq d+1$.*

Proof. For every basis a we have that $\sum_{i=1}^{d} |\Phi_a(i)\rangle\langle\Phi_a(i)| = \mathbb{1}$. Hence, every basis leads to at most $(d-1)$ independent projectors and, together with the identity, the bases span an at most $(k(d-1)+1)$-dimensional space. To be tomographically complete we require that

$$k(d-1)+1 \geq d^2.$$

Therefore we need $k \geq (d+1)$. ■

So in the case $k = d+1$, we can compare solutions for mutually unbiased bases with solutions for tomographically complete bases.

6.2 Strategies and Marginals of Joint Distributions

In order to survive, Alice has to choose her initial state ρ, her measurements F_x and her estimation function s such that (see Figure 6.1)

$$\operatorname{tr}\big[\rho(|\Phi_a(i)\rangle\langle\Phi_a(i)| \otimes \mathbb{1})F_x(|\Phi_a(i)\rangle\langle\Phi_a(i)| \otimes \mathbb{1})\big] = \begin{cases} \lambda_{ix} & \text{if } i = s(x,a), \\ 0 & \text{otherwise.} \end{cases} \tag{6.3}$$

Here, $\lambda_{ix} \geq 0$ such that $\sum_i \sum_x \lambda_{ix} = 1$. That is, Alice guess is never wrong and the total probability of getting a result given the combination of measurements $\{\Phi_a(i)\}$ and $\{F_x\}$ is one. To simplify the notation, we take the measurement result x to be itself the mapping $x\colon a \mapsto i$. Then equation (6.3) becomes

$$\operatorname{tr}\big[\rho(|\Phi_a(i)\rangle\langle\Phi_a(i)| \otimes \mathbb{1})F_x(|\Phi_a(i)\rangle\langle\Phi_a(i)| \otimes \mathbb{1})\big] = \lambda_{ix}\delta_{i,x(a)}. \tag{6.4}$$

As there are k possible choices for the basis and d different measurement results, the number of possible estimation functions or equivalently the number of different outcomes $|X|$, $X = \big\{x\colon \{1,\ldots,k\} \to \{1,\ldots,d\}\big\}$, is $|X| = d^k$. We will call Alice choice of initial state ρ and measurement $\{F_x\}$ a strategy if they satisfy (6.4) for the given bases $\{\Phi_a(i)\}$.

Intuitively, Alice will try to keep a system that is maximally correlated with the system she gave to the king. This suggests that a maximally entangled state is the optimal choice for an initial state.

First, note that Alice can always choose a pure state as initial state, even if we do not pose any constraint on the king's bases.

6.2.1 Lemma. *If Alice found a solution for equation* (6.4) *with initial state ρ, then there exists a solution with a pure initial state.*

Proof. Assume that ρ is a solution of (6.4). Let $\rho = \sum_j \lambda_j |e_j\rangle\langle e_j|$, $\lambda_j > 0$, be a decomposition of ρ into normalized pure states $|e_j\rangle$. Since $\lambda_j > 0$ we have

$$\langle e_j|(|\Phi_a(i)\rangle\langle\Phi_a(i)|\otimes\mathbb{1})F_x(|\Phi_a(i)\rangle\langle\Phi_a(i)|\otimes\mathbb{1})|e_j\rangle = 0$$

for all $i \neq x(a)$ and all j. Also, as $0 < \lambda_j \leq 1$, $\sum_j \lambda_j = 1$, and

$$(|\Phi_a(i)\rangle\langle\Phi_a(i)|\otimes\mathbb{1})F_x(|\Phi_a(i)\rangle\langle\Phi_a(i)|\otimes\mathbb{1}) \leq \mathbb{1},$$

we have

$$\langle e_j|(|\Phi_a(i)\rangle\langle\Phi_a(i)|\otimes\mathbb{1})F_x(|\Phi_a(i)\rangle\langle\Phi_a(i)|\otimes\mathbb{1})|e_j\rangle = 1$$

for all $i = x(a)$ and all j. Hence $\rho = |e_1\rangle\langle e_1|$ is a solution to (6.4) with an initial pure state. ■

Suppose Alice has found a solution to (6.4) and the corresponding pure initial state of Lemma 6.2.1 is $|\Psi\rangle \in \mathcal{H}\otimes\mathcal{K}$, $\dim\mathcal{H} = d$, $\dim\mathcal{K} = D$. Let $|\Omega\rangle = 1/\sqrt{d}\sum_i |ii\rangle$ be the maximally entangled state on $\mathcal{H}\otimes\mathcal{H}$. Due to the Schmidt decomposition of $|\Psi\rangle$, $|\Psi\rangle = \sum_j \lambda_j |j\rangle_{\mathcal{H}}|j\rangle_{\mathcal{K}}$, we can already assume that $\dim\mathcal{K} = \dim\mathcal{H}$. However, we write $|\Psi\rangle$ as

$$|\Psi\rangle = (\mathbb{1}\otimes S)|\Omega\rangle = \frac{1}{\sqrt{d}}\sum_{i=1}^{D}\sum_{j=1}^{d}\sum_{k=1}^{d}(\mathbb{1}\otimes|i\rangle\langle j|)S_{ij}|k\rangle|k\rangle = \sum_{i=1}^{D}\sum_{j=1}^{d}\frac{S_{ij}}{\sqrt{d}}|j\rangle|i\rangle,$$

with an operator $S\colon\mathcal{H}\to\mathcal{K}$. As $(\mathbb{1}\otimes S)$ commutes with the king's men's measurement and with $|\hat{\Phi}_a(i)\rangle \in \mathcal{H}\otimes\mathcal{H}$ defined as

$$\begin{aligned}|\hat{\Phi}_a(i)\rangle :&= (|\Phi_a(i)\rangle\langle\Phi_a(i)|\otimes\mathbb{1})|\Omega\rangle \\ &= \frac{1}{\sqrt{d}}\sum_{\alpha}\langle\Phi_a(i)|\alpha\rangle|\Phi_a(i)\rangle|\alpha\rangle = \frac{1}{\sqrt{d}}|\Phi_a(i)\rangle|\overline{\Phi_a(i)}\rangle,\end{aligned} \tag{6.5}$$

condition (6.4) becomes

$$\langle\hat{\Phi}_a(i)|(\mathbb{1}\otimes S)^* F_x(\mathbb{1}\otimes S)|\hat{\Phi}_a(i)\rangle = \lambda_{ix}\delta_{i,x(a)}. \tag{6.6}$$

Alice strategy would stay the same if she uses a maximally entangled state initially and if she could measure

$$\hat{F}_x = (\mathbb{1}\otimes S^*)F_x(\mathbb{1}\otimes S) \tag{6.7}$$

instead of F_x. Clearly $\hat{F}_x$ are positive as F_x are positive. It remains to show that $\sum_x \hat{F}_x = \mathbb{1}$ in order to be a valid measurement. In this case, Alice could start her strategy with a maximally entangled state.

Now we look at the structure of Alice strategy. Again, assume Alice has a strategy and let $|\eta_x\rangle$ be a vector in the support of the operator $\hat{F}_x$. From (6.6) it follows that

$$\langle\hat{\Phi}_a(i)||\eta_x\rangle\langle\eta_x||\hat{\Phi}_a(i)\rangle = 0$$

whenever $i \neq x(a)$, and hence we have

$$\langle\hat{\Phi}_a(i)|\eta_x\rangle = 0 \quad \text{if} \quad i \neq x(a). \tag{6.8}$$

Definition (6.5) implies that $\sum_i |\hat{\Phi}_a(i)\rangle = |\Omega\rangle$. Therefore, we get

$$\langle\hat{\Phi}_a(i)|\eta_x\rangle = \langle\hat{\Phi}_a(i)|\eta_x\rangle + \sum_{j\neq x(a)} \langle\hat{\Phi}_a(j)|\eta_x\rangle = \langle\Omega|\eta_x\rangle\delta_{i,x(a)}. \tag{6.9}$$

To determine how many of such vectors $|\eta_x\rangle$ Alice can find, we look at the space $\mathcal{R}$ defined in (6.1), or equivalently at the space

$$\hat{\mathcal{R}} := \mathrm{span}_{\mathbb{R}}\{|\hat{\Phi}_a(i)\rangle\}$$

by the identification (6.5). If Alice can't find any vector satisfying (6.9) with $\langle\Omega|\eta_x\rangle \neq 0$, then all solutions of that equation are in the orthogonal complement $\hat{\mathcal{R}}^\perp$ of $\hat{\mathcal{R}}$. This implies that the value x can never occur as a result of Alice's measurement. Thus, Alice strategy has to rely on the cases where such solutions exist. In the following, we divide such a solution by $\langle\Omega|\eta_x\rangle$ to have

$$\langle\hat{\Phi}_a(i)|\eta_x\rangle = \delta_{i,x(a)}. \tag{6.10}$$

Equation (6.10) is a linear system of equations with complex entries. If we restrict $|\eta_x\rangle$ to be a vector in $\hat{\mathcal{R}}_{\mathbb{C}} := \mathrm{span}_{\mathbb{C}}\{|\hat{\Phi}_a(i)\rangle\}$, the solution is uniquely determined, since all scalar products with vectors from this space are fixed. Furthermore, we then have $|\eta_x\rangle \in \hat{\mathcal{R}}$, as all scalar products are real. This can immediately be seen if we write $|\eta_x\rangle \in \hat{\mathcal{R}}_{\mathbb{C}}$ as $|\eta_x\rangle = |\alpha_x\rangle + i|\beta_x\rangle$ with $|\alpha_x\rangle, |\beta_x\rangle \in \hat{\mathcal{R}}$. Consequently, whenever a non-zero solution exists for some $x \in X$, we can pick a unique solution η_x, determined by the conditions

$$\boxed{\eta_x \in \hat{\mathcal{R}}, \quad \langle\hat{\Phi}_a(i)|\eta_x\rangle = \delta_{i,x(a)}.} \tag{6.11}$$

Remember that vectors in $\hat{\mathcal{R}}$ correspond to hermitian operators. Expressed in the standard basis we have

$$\sqrt{d}\sum_{\alpha,\beta} |\alpha\rangle\langle\beta|\langle\alpha\beta|\eta_x\rangle = \left(\sqrt{d}\sum_{\alpha,\beta} |\alpha\rangle\langle\beta|\langle\alpha\beta|\eta_x\rangle\right)^*$$
$$= \sqrt{d}\sum_{\alpha,\beta} |\beta\rangle\langle\alpha|\overline{\langle\alpha\beta|\eta_x\rangle} = \sqrt{d}\sum_{\alpha,\beta} |\beta\rangle\langle\alpha|\langle\beta\alpha|\eta_x\rangle$$

and hence

$$\langle \alpha\beta|\eta_x\rangle = \overline{\langle \beta\alpha|\eta_x\rangle}. \tag{6.12}$$

It follows from the above that the measurement of an effect proportional to $|\eta_x\rangle\langle\eta_x|$ would have the result "true" only if $i = x(a)$, and "false" otherwise.

6.2.2 Proposition. *If Alice can find any vector $|\eta_x\rangle$ satisfying* (6.11)*, then a measurement $\{\lambda|\eta_x\rangle\langle\eta_x|, \mathbb{1} - \lambda|\eta_x\rangle\langle\eta_x|\}$, $\lambda = 1/\|\eta_x\|^2$, would be a strategy for an unambiguous retrodiction problem, where she is allowed to pass (measurement of $\mathbb{1} - \lambda|\eta_x\rangle\langle\eta_x|$), but has to be absolutely sure of her guess otherwise (measurement of $\lambda|\eta_x\rangle\langle\eta_x|$).*

That Alice has to be right with her guess in every run, and in particular independent of the king's result i, therefore imposes a non-trivial constraint: Alice has to find $\lambda_x \geq 0$ and η_x satisfying (6.11), such that $\sum_x \lambda_x|\eta_x\rangle\langle\eta_x| = \mathbb{1}$.

We will now show that strategies for Alice are related to probability distributions over X.

6.2.3 Definition (Classical Model)**.** A set of k bases $\{|\Phi_a(i)\rangle\}$ admits a classical model, if there exists a joint distribution p of k variables $l_i \in \{1, \ldots, d\}$ with probability mass function $p(l_1, \ldots, l_k)$, such that the marginals

$$p_{ab}(i,j) := \sum_{l_1,\ldots,l_k=1}^{d} \delta_{l_a,i}\delta_{l_b,j}\, p(l_1, \ldots, l_k)$$

are given by

$$p_{ab}(i,j) = \frac{1}{d}|\langle \Phi_a(i)|\Phi_b(j)\rangle|^2. \tag{6.13}$$

6.2.4 Theorem. *Let $\{|\Phi_a(i)\rangle\}$ be a collection of k orthonormal bases in a d-dimensional Hilbert space. Then we have:*

1. *If the bases are non-degenerate and admit a classical model, then there exists a strategy for Alice that solves the meaner king problem and Alice can use a maximally entangled state initially.*

2. *If the set of bases is tomographically complete and Alice has a strategy that solves the meaner king problem, then the king's bases admit a classical model and if Alice's strategy begins with a pure state, this state is maximally entangled.*

In particular we have $k \leq (d+1)$ in the first case and $k \geq (d+1)$ in the second case. Note that the existence of a classical model does only depend on the absolute value of the scalar products (6.13) and not on their phases. That is, only part of the available information about the bases is actually required for the theorem.

Proof of Theorem 6.2.4. Due to Lemma 6.2.1, we will only consider strategies with pure initial states.

The first part of the theorem states sufficient conditions for the existence of a strategy for Alice. Suppose that the bases are non-degenerate and that a classical model exists. For each x, condition (6.11) is a linear system of equations for $|\eta_x\rangle$. As $\sum_i |\hat{\Phi}_a(i)\rangle = |\Omega\rangle$ for every basis, we only take the first $(d-1)$ equations (6.11) for each a. Together with the normalization condition $\langle\Omega|\eta_x\rangle = 1$, we have a system of $k(d-1)+1$ equations. Non-degeneracy of the bases guarantees that these equations are non-singular, and therefore one can find a vector $|\eta_x\rangle$ for every x. Thus we have

$$\begin{aligned}\langle\hat{\Phi}_a(i)|\sum_x p(x)|\eta_x\rangle\langle\eta_x||\hat{\Phi}_b(j)\rangle &= \sum_x p(x)\delta_{i,x(a)}\delta_{j,x(b)} \\ &= \frac{1}{d}|\langle\Phi_a(i)|\Phi_b(j)\rangle|^2 = \langle\hat{\Phi}_a(i)|\hat{\Phi}_b(j)\rangle,\end{aligned} \tag{6.14}$$

where we used that fact that p is a classical model in the second equation. This amounts to

$$\sum_x p(x)|\eta_x\rangle\langle\eta_x| = \mathbb{1}_{\hat{\mathcal{R}}}. \tag{6.15}$$

Furthermore, we can add to each operator $\Gamma_x = p(x)|\eta_x\rangle\langle\eta_x|$ an operator $\Gamma_x^\perp \geq 0$ from the complement of $\hat{\mathcal{R}}$, such that $\sum_x(\Gamma_x + \Gamma_x^\perp) = \mathbb{1}\otimes S^*S$ for some S with $\|(\mathbb{1}\otimes S)|\Omega\rangle\| = 1$. This provides a valid strategy for Alice, since the operators $\Gamma_x^\perp$ have no influence on the measured expectation values. In particular, we can choose $S = \mathbb{1}$, i. e., $\sum_x \Gamma_x^\perp$ is the projection onto $\hat{\mathcal{R}}^\perp$, which means that Alice can use a maximally entangled state initially.

Now, we want to prove the second part of the theorem, i. e., show the necessary conditions for the existence of a strategy in the case that the bases are tomographically complete. Suppose Alice has a strategy and the bases are tomographically complete. As the bases are tomographically complete, we have $\hat{\mathcal{R}}^\perp = \{0\}$. Therefore, Alice strategy has to consist of vectors according to (6.11) with scalar factors $p(x) \geq 0$,

$$\hat{F}_x = p(x)|\eta_x\rangle\langle\eta_x|.$$

Note that we can set $p(x) = 0$ for any x for which (6.11) has no non-zero solution to have $\operatorname{dom} p = X$. Since $p(x) \geq 0$, we have $\hat{F}_x \geq 0$. If we expand $|\eta_x\rangle$ in the standard basis,

$$|\eta_x\rangle = \sum_{\alpha,\beta}\langle\alpha\beta|\eta_x\rangle|\alpha\beta\rangle,$$

the overall normalization condition becomes

$$\begin{aligned}(\mathbb{1}\otimes S^*S) = \sum_x \hat{F}_x &= \sum_x p(x)|\eta_x\rangle\langle\eta_x| \\ &= \sum_{\alpha,\beta}\sum_{\gamma,\delta}\sum_x p(x)\langle\alpha\beta|\eta_x\rangle|\alpha\beta\rangle\langle\gamma\delta|\overline{\langle\gamma\delta|\eta_x\rangle}.\end{aligned}$$

Using (6.12) and since the first tensor factor is unity, we know that

$$\begin{aligned}\sum_x p(x)\langle\alpha\beta|\eta_x\rangle\langle\delta\gamma|\eta_x\rangle &= \delta_{\alpha\gamma}u_{\beta\delta}\\ = \sum_x p(x)\overline{\langle\beta\alpha|\eta_x\rangle\langle\gamma\delta|\eta_x\rangle} &= \delta_{\beta\delta}\overline{u_{\alpha\gamma}},\end{aligned}$$

for some $u_{\beta\delta} \in \mathbb{C}$. Furthermore, since $S^*S \geq 0$, we know that $u_{\beta\delta} \geq 0$. In particular, if we set $\gamma = \alpha$ we get $u_{\beta\delta} = \delta_{\beta\delta}u_{\alpha\alpha}$ independently of α. Since $1 = \|\Psi\|^2 = \langle\Omega|\mathbb{1}\otimes S^*S|\Omega\rangle = u_{\alpha\alpha}$, we conclude that $u_{\alpha\alpha} = 1$. Thus we have

$$\sum_x p(x)\langle\alpha\beta|\eta_x\rangle\langle\delta\gamma|\eta_x\rangle = \delta_{\alpha\gamma}\delta_{\beta\delta},$$

which amounts to $S^*S = \mathbb{1}$: in the tomographically complete case, the initial state of Alice must be maximally entangled.

Now consider the scalar product of $\sum_x \hat{F}_x$ with $|\Omega\rangle$,

$$1 = \langle\Omega|\Omega\rangle = \langle\Omega|\sum_x \hat{F}_x|\Omega\rangle = \sum_x p(x)|\langle\Omega|\eta_x\rangle|^2 = \sum_x p(x),$$

where we used (6.11) and $S^*S = \mathbb{1}$. Since we already know that $p(x) \geq 0$, we obtain that p is a probability distribution over X. On the other hand, if we take scalar products with $\hat{\Phi}_a(i)$ as defined in (6.5) and use (6.11), we get

$$\begin{aligned}\frac{1}{d}|\langle\Phi_a(i)|\Phi_b(j)\rangle|^2 &= \langle\hat{\Phi}_a(i)|\hat{\Phi}_b(j)\rangle\\ = \langle\hat{\Phi}_a(i)|\sum_x p(x)|\eta_x\rangle\langle\eta_x||\hat{\Phi}_b(j)\rangle &= \sum_x p(x)\delta_{i,x(a)}\delta_{j,x(b)}.\end{aligned}$$

Hence the king's bases admit the classical model p according to Definition 6.2.3. ■

6.2.5 Corollary. *Let $\{|\Phi_b(i)\rangle\}$ be a set of k mutually unbiased bases in a d-dimensional Hilbert space, then Alice can find a strategy for the mean king problem 6.0.2.*

Proof. We have already seen in Lemma 6.1.5 that mutually unbiased bases are non-degenerate. Furthermore, the marginals are given by

$$p_{ab}(i,j) = \frac{1}{d}|\langle\Phi_a(i)|\Phi_b(j)\rangle|^2 = \delta_{ab}\delta_{ij}\frac{1}{d} + (1-\delta_{ab})\frac{1}{d^2}, \tag{6.16}$$

where we used the mutually unbiasedness condition in the second equality. These are exactly the marginals of k statistical independent uniformly distributed random variables. Thus, a safe strategy for Alice exists according to Theorem 6.2.4. ■

6.3 Finding a Strategy for Alice

Given the bases of the king, we want to determine whether we can find a strategy for Alice. We will assume that the given bases are non-degenerate unless stated otherwise, that is, we are considering the meaner king problem as described in Figure 6.2. According to Theorem 6.2.4, a strategy for Alice exists, if we can find a joint probability p with the marginals

$$p_{ab}(i,j) = \frac{1}{d}|\langle\Phi_a(i)|\Phi_b(j)\rangle|^2, \tag{6.17}$$

for $a, b = 1, \ldots, k$ and $i, j = 1, \ldots, d$. Observe that for any pair (a, b) of bases, the values $p_{ab}(i, j)$ are the joint probabilities of a pair of d-valued, uniformly distributed random variables:

$$\begin{aligned}\sum_i p_{ab}(i,j) &= \frac{1}{d}\langle\Phi_b(j)|\sum_i |\Phi_a(i)\rangle\langle\Phi_a(i)||\Phi_b(j)\rangle = \frac{1}{d} \\ &= \frac{1}{d}\langle\Phi_a(i)|\sum_j |\Phi_b(j)\rangle\langle\Phi_b(j)||\Phi_a(i)\rangle = \sum_j p_{ab}(i,j)\end{aligned} \tag{6.18}$$

and

$$\sum_i \sum_j p_{ab}(i,j) = \sum_{i=1}^{d} \frac{1}{d} = 1. \tag{6.19}$$

An particularly interesting case is the setting $d = 6$ and $k = 7$, as it is still an open problem, whether mutually unbiased bases exist in this case.

6.3.1 Qubit case

The simplest non-trivial case is $d = 2$ and $k = 3$. One choice for non-degenerate bases would be the eigenvectors of the three Pauli matrices, as they are mutually unbiased. In general, a basis in $d = 2$ is given by the two intersecting points of a line through the origin of the Bloch sphere with the sphere's hull. Three qubit bases are non-degenerate (and hence tomographically complete), if these lines do not lie in a plane. For special choices of these bases, the problem has been discussed by Ben-Menahem [113].

6.3.1.1 Possible Situations

In the qubit case, the marginals (6.17) are determined by three parameters $p_{12}, p_{13}, p_{23} \in [0, 1]$,

$$p_{12} := p_{12}(1,1), \quad p_{13} := p_{13}(1,1), \quad p_{23} := p_{23}(1,1). \tag{6.20}$$

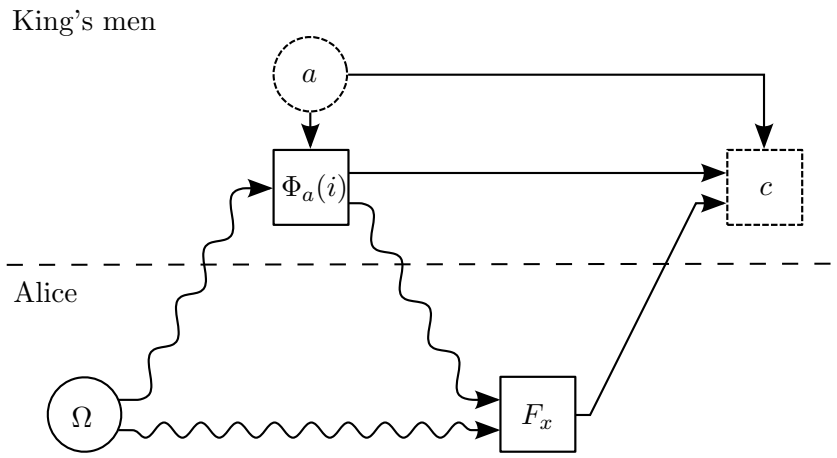

Figure 6.2: The meaner king problem in the non-degenerate case: The king's men choose between k non-degenerate bases $\{\Phi_a(i)\}$, $i = 1, \ldots, d$. They make a von Neumann measurement on their part of the maximally entangled state Ω but keep their result i secret. Afterwards, Alice does a measurement F_x on both systems and sends her estimation function $x\colon a \mapsto i$ to the king. The results are compared (function c) and Alice dies a cruel death if $x(a) \neq i$.

All other marginals follow from the equations (6.18) and (6.19):

$$p_{ab}(2,1) = \frac{1}{2} - p_{ab}(1,1), \quad p_{ab}(1,2) = \frac{1}{2} - p_{ab}(1,1), \tag{6.21}$$

$$p_{ab}(2,2) = 1 - p_{ab}(1,1) - p_{ab}(2,1) - p_{ab}(1,2) = p_{ab}(1,1), \tag{6.22}$$

where $(a,b) = (1,2), (1,3), (2,3)$.

These parameters are not independent. Suppose p_{12} and p_{13} are given, then we have

$$p_{23} = \frac{1}{2}|\langle\Phi_2(1)|\Phi_3(1)\rangle|^2.$$

We expand $|\Phi_2(1)\rangle$ in the basis $|\Phi_1(i)\rangle$ using the equations (6.17) and (6.21) and obtain:

$$\begin{aligned}
|\Phi_2(1)\rangle &= \sum_{j=1}^{2} \langle\Phi_1(j)|\Phi_2(1)\rangle|\Phi_1(j)\rangle \\
&= \langle\Phi_1(1)|\Phi_2(1)\rangle|\Phi_1(1)\rangle + \langle\Phi_1(2)|\Phi_2(1)\rangle|\Phi_1(2)\rangle \\
&= e^{i\alpha}\sqrt{2\,p_{12}}|\Phi_1(1)\rangle + e^{i\beta}\sqrt{1-2\,p_{12}}|\Phi_1(2)\rangle,
\end{aligned}$$

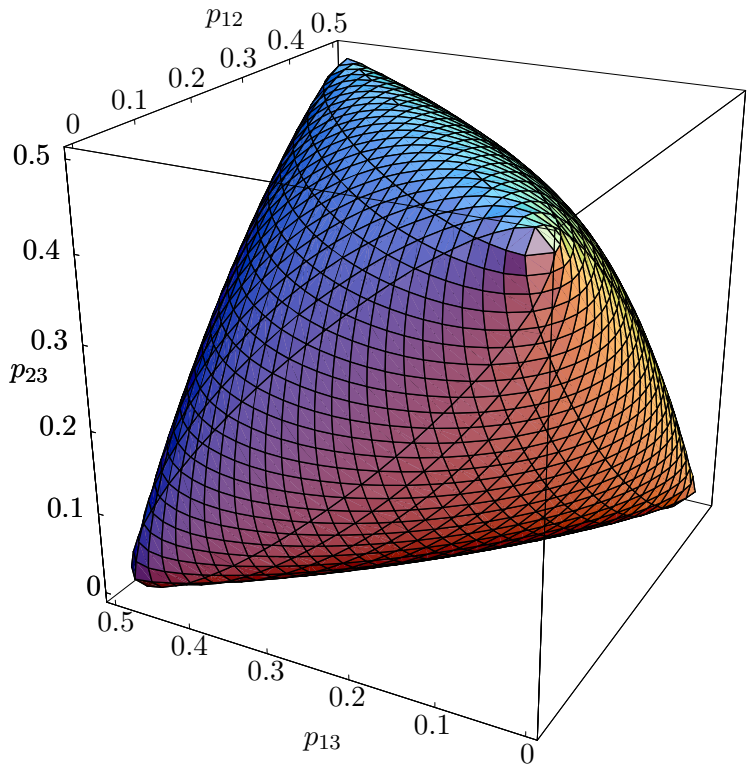

Figure 6.3: Possible range of marginals (p_{12}, p_{13}, p_{23}) for the qubit case of the meaner king problem. The range of triples admitting a classical model and therefore a successful retrodiction strategy for Alice is described as tetrahedron inside this body.

for some phases $\alpha, \beta \in \mathbb{R}$. Then, p_{23} reads

$$\begin{aligned} 2\,p_{23} &= |e^{-i\alpha}\sqrt{2\,p_{12}}\langle\Phi_1(1)|\Phi_3(1)\rangle + e^{-i\beta}\sqrt{1-2\,p_{12}}\langle\Phi_1(2)|\Phi_3(1)\rangle|^2 \\ &= |e^{i\gamma}\sqrt{2\,p_{12}2\,p_{13}} + e^{i\vartheta}\sqrt{(1-2\,p_{12})(1-2\,p_{13})}|^2 \\ &= |\sqrt{2\,p_{12}2\,p_{13}} + e^{i\varphi}\sqrt{(1-2\,p_{12})(1-2\,p_{13})}|^2 \\ &= 2\,p_{12}d\,p_{13} + (1-2\,p_{12})(1-2\,p_{13}) \\ &\quad + 2\cos(\varphi)\sqrt{2\,p_{12}2\,p_{13}(1-2\,p_{12})(1-2\,p_{13})}, \end{aligned}$$

for some phases $\gamma, \vartheta, \varphi \in \mathbb{R}$. Thus, the value of p_{23} is bracketed by the cases $\cos(\varphi) = \pm 1$ as shown in Figure 6.3.

6.3.1.2 Situations with a Solution

We will now show that the marginals admitting a classical model correspond to a tetrahedron inside the baggy tetrahedron in Figure 6.3. To this end, we consider the original Bell setting [20]. The outcome of the measurement of basis a corresponds

to a random variable x_a with the values $\{1, 2\}$. Let $x = (x_1, x_2, x_3)$ be the vector of random variables for the three bases. We assign a sign to the value x_a by

$$\sigma_a(x) := \begin{cases} +1 & \text{if } x_a = 1, \\ -1 & \text{if } x_a = 2. \end{cases} \tag{6.23}$$

That is, $\sigma_a(x)$ is $+1$, if the outcome of the random variable for the a-th basis is 1 and -1 otherwise. As in Bell's setting, we define the correlation coefficient to be

$$C_{ab} := \sum_{x_1=1}^{2} \sum_{x_2=1}^{2} \sum_{x_3=1}^{2} p(x)\sigma_a(x)\sigma_b(x). \tag{6.24}$$

Here $p(x)$ is the probability mass function of a classical probability distribution and a and b are the indices of the random variables. For given marginals $p_{ab}(i, j)$, equation (6.24) becomes

$$C_{ab} := \sum_{x_a=1}^{2} \sum_{x_b=1}^{2} p_{ab}(x_a, x_b)\sigma(x_a)\sigma(x_b), \tag{6.25}$$

where σ is defined analogously to (6.23). If we look at the difference between two correlation coefficients, we get

$$\begin{aligned} C_{ab} - C_{ac} &= \sum_x p(x)(\sigma_a(x)\sigma_b(x) - \sigma_a(x)\sigma_c(x)) \\ &= \sum_x p(x)\sigma_a(x)\sigma_b(x)\left(1 - \sigma_b(x)\sigma_c(x)\right). \end{aligned}$$

Here we used the fact that $\sigma_b(x)^2 = 1$ in the second equation. Furthermore, $\big(1 - \sigma_b(x)\sigma_c(x)\big) \geq 0$, so we can conclude that

$$C_{ab} - C_{ac} \leq \sum_x p(x)\left(1 - \sigma_b(x)\sigma_c(x)\right) = 1 - C_{bc}. \tag{6.26}$$

Together with the inequalities that follow from the substitutions $\sigma_a(x) \to -\sigma_a(x)$, $\sigma_b(x) \to -\sigma_b(x)$, and $\sigma_c(x) \to -\sigma_c(x)$, we obtain the four Bell inequalities

$$C_{ab} - C_{ac} + C_{bc} \leq 1 \qquad \text{original eq. (6.26)}, \tag{6.27}$$

$$-C_{ab} + C_{ac} + C_{bc} \leq 1 \qquad \text{with } \sigma_a(x) \to -\sigma_a(x), \tag{6.28}$$

$$-C_{ab} - C_{ac} - C_{bc} \leq 1 \qquad \text{with } \sigma_b(x) \to -\sigma_b(x), \tag{6.29}$$

$$C_{ab} + C_{ac} - C_{bc} \leq 1 \qquad \text{with } \sigma_c(x) \to -\sigma_c(x). \tag{6.30}$$

We now express the correlations C_{ab} in terms of the parameters p_{ab} defined in (6.20). From equation (6.25), we get

$$C_{ab} = p_{ab}\sigma(1)^2 + 2\left(\frac{1}{2} - p_{ab}\right)\sigma(1)\sigma(2) + p_{ab}\sigma(2)^2 = 4p_{ab} - 1,$$

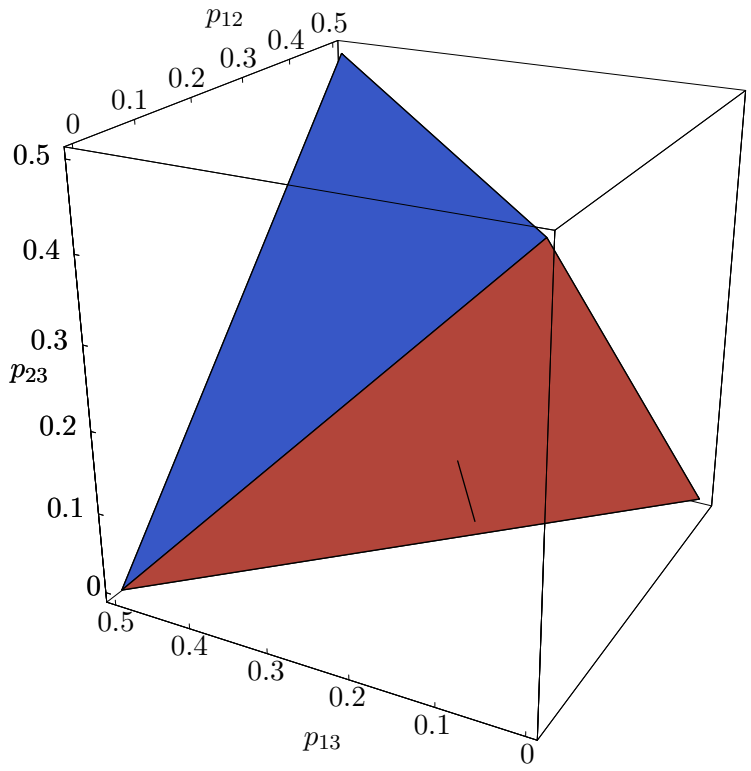

Figure 6.4: Possible range of marginals (p_{12}, p_{13}, p_{23}) admitting a classical model. This tetrahedron is described by the inequalities (6.31) and lies inside the baggy tetrahedron depicted in Figure 6.3. The black line shows the marginals (6.34) of the parametrized bases (6.33) for the range (6.35) of the parameter θ.

where we used (6.21) and (6.22). With the choice of parameters $p = 2p_{12}$, $q = 2p_{13}$ and $r = 2p_{23}$, the inequalities (6.27), (6.28), (6.29), and (6.30) describe the four sides of a tetrahedron inside the baggy tetrahedron in Figure 6.3,

$$\begin{aligned}
\text{eq. (6.27)} &\Leftrightarrow \quad (2p-1)-(2q-1)+(2r-1)\leq 1 \Leftrightarrow \quad p-q+r\leq 1,\\
\text{eq. (6.28)} &\Leftrightarrow \quad -(2p-1)+(2q-1)+(2r-1)\leq 1 \Leftrightarrow \quad -p+q+r\leq 1,\\
\text{eq. (6.29)} &\Leftrightarrow \quad -(2p-1)-(2q-1)-(2r-1)\leq 1 \Leftrightarrow \quad p+q+r\geq 1,\\
\text{eq. (6.30)} &\Leftrightarrow \quad (2p-1)+(2q-1)-(2r-1)\leq 1 \Leftrightarrow \quad p+q-r\leq 1.
\end{aligned} \tag{6.31}$$

This tetrahedron is depicted in Figure 6.4. Hence, a classical model for given marginals only exists, if these marginals correspond to a point inside that tetrahedron. Therefore, even in the qubit-case, there are quantum mechanical values of $p_{ab}(i,j)$ that do not correspond to a classical model. These values correspond to the bulge regions in Figure 6.3. As a consequence, if the king can choose tomographically complete bases with marginals outside the classical tetrahedron, Alice cannot solve the retrodiction problem with certainty, according to Theorem 6.2.4.

6.3.1.3 A Mean Choice

We will now give an example of such a choice of bases, where Alice cannot find a strategy to retrodict the king's outcome with certainty. To this end, we parametrize the pure states with angles θ and φ as shown in Figure 6.5,

$$|\Psi(\theta,\varphi)\rangle = \cos\frac{\theta}{2}|0\rangle + e^{i\varphi}\sin\frac{\theta}{2}|1\rangle, \tag{6.32}$$

where $|0\rangle$ and $|1\rangle$ are the eigenstates of the σ_z Pauli operator for the positive and negative eigenvalue, respectively. For a given angle θ, we choose the vectors of the three bases to be

$$\begin{aligned}
\{\Psi_1(1),\Psi_1(2)\} &= \{|\Psi(\theta,0)\rangle, |\Psi(\theta+\pi,0)\rangle\},\\
\{\Psi_2(1),\Psi_2(2)\} &= \{|\Psi(\theta,\tfrac{2\pi}{3})\rangle, |\Psi(\theta+\pi,\tfrac{2\pi}{3})\rangle\},\\
\{\Psi_3(1),\Psi_3(2)\} &= \{|\Psi(\theta,\tfrac{4\pi}{3})\rangle, |\Psi(\theta+\pi,\tfrac{4\pi}{3})\rangle\}.
\end{aligned} \tag{6.33}$$

Clearly, as $\cos(\alpha+\pi/2) = -\sin(\alpha)$ and $\sin(\alpha+\pi/2) = \cos(\alpha)$, these are three orthogonal bases. Furthermore, for $0 < \theta < \pi/2$, the Bloch vectors of these bases do not lie in a plane and the bases are therefore tomographically complete.

By construction of the bases (6.33), all parameters p_{12}, p_{13}, and p_{23} are equal. This can be seen in Figure 6.5, but also through direct calculation: Given the vectors $|\Psi_a(1)\rangle = |\Psi(\theta,\varphi)\rangle$ and $|\Psi_b(1)\rangle = |\Psi(\theta,\omega)\rangle$, the corresponding marginal $p_{ab}(1,1)$ is given by

$$\begin{aligned}
p_{ab}(1,1) &= \frac{1}{2}|\langle\Psi(\theta,\varphi)|\Psi(\theta,\omega)\rangle|^2\\
&= \frac{1}{2}\left|\cos\left(\frac{\theta}{2}\right)^2 + e^{i(\omega-\varphi)}\sin\left(\frac{\theta}{2}\right)^2\right|^2\\
&= \frac{1}{2}\left(\cos\left(\frac{\theta}{2}\right)^4 + \sin\left(\frac{\theta}{2}\right)^4 + 2\cos\left(\frac{\theta}{2}\right)^2\sin\left(\frac{\theta}{2}\right)^2\cos(\omega-\varphi)\right)\\
&= \frac{1}{2}\left(1 + \frac{\sin(\theta)^2}{2}\left(\cos(\omega-\varphi)-1\right)\right).
\end{aligned}$$

For p_{12} and p_{23} we have $(\omega-\varphi) = 2\pi/3$ and for p_{13} we have $(\omega-\varphi) = 4\pi/3$. But since $\cos 2\pi/3 = \cos 4\pi/3 = -1/2$, all values are equal to

$$p_{12} = p_{13} = p_{23} = \frac{1}{2}\left(1 - \frac{3}{4}\sin(\theta)^2\right) = \frac{1}{16}\left(3\cos(2\theta)+5\right). \tag{6.34}$$

Consequently, the values of (6.34) lie between $1/2$ for $\theta = 0$ and $1/8$ for $\theta = \pi/2$. In the Figures 6.3 and 6.4, this corresponds to the diagonal line segment from $p_{12} = p_{13} = p_{23} = 1/2$ to $p_{12} = p_{13} = p_{23} = 1/8$. The line segment intersects the face $p_{12}+p_{13}+p_{23} \geq 1/2$ of the tetrahedron (6.31) at $\theta = 1/2\arccos(-7/9)$.

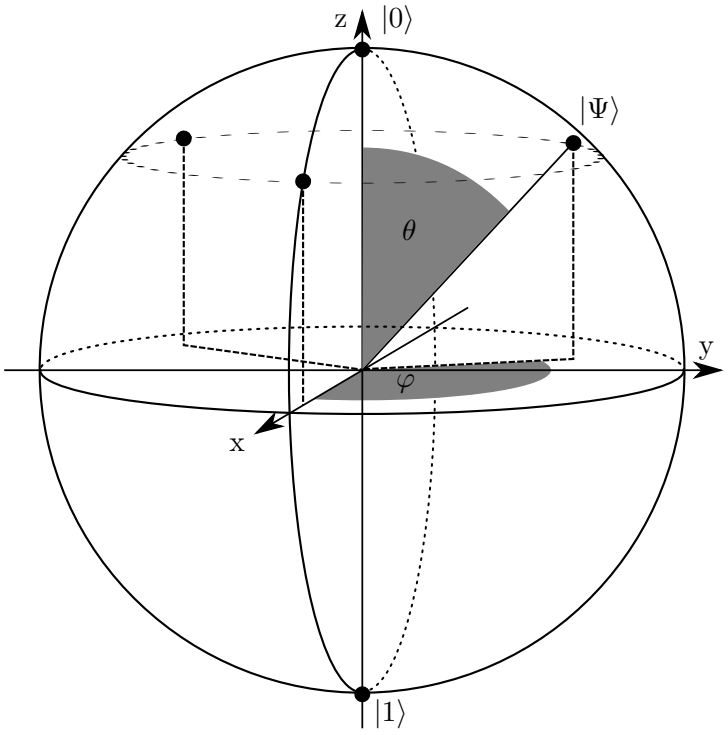

Figure 6.5: Bloch sphere representations of the choice of the first basis vector for the three bases. The vectors have the same angle θ and the angles $\varphi = 0$, $\varphi = 2\pi/3$ and $\varphi = 4\pi/3$.

Thus, for

$$\frac{1}{2}\arccos\left(-\frac{7}{9}\right) < \theta < \frac{\pi}{2}, \tag{6.35}$$

the bases (6.33) are tomographically complete and do not admit a classical model. Therefore, Theorem 6.2.4 tells us that Alice cannot find a strategy for the meaner king problem 6.1.1 that guarantees her to successfully retrodict the king's outcome in every case. This shows that, even in the qubit case, the situation is dramatically different to the mean king problem with mutually unbiased bases (Problem 6.0.2), where Alice can always find a successful strategy.

6.3.1.4 A Random Choice for the King's Bases

We are now going to answer what Alice chance is to find a strategy, if the king chooses his bases randomly according to the Haar measure. The basis vectors of an orthonormal bases can be seen as the columns of a unitary matrix, so the king's qubit bases are elements of the unitary group $U(2)$. The Haar measure is the unique

measure μ, $\mu(U(2)) = 1$, that is left-translational invariant in that group, $\mu(\Gamma) = \mu(\{Ug\,|g \in \Gamma\})$ for all $U \in U(2)$, and Γ in the Borel subsets of $U(2)$. This means that the probability for the bases are the same, independently of an additional rotation of the Bloch sphere.

As we are aware of the region of triples (p_{12}, p_{13}, p_{23}) for which a classical model and hence a strategy for Alice exist, we have to compute the probability that this triple is outside the tetrahedron (6.31) delimiting this region. First, let us compute the probability to be outside the face $p_{12} + p_{13} + p_{23} \geq 1/2$, i. e., the probability for

$$2(p_{12}+p_{13}+p_{23}) = |\langle\Phi_1(1)|\Phi_2(1)\rangle|^2+|\langle\Phi_1(1)|\Phi_3(1)\rangle|^2+|\langle\Phi_2(1)|\Phi_3(1)\rangle|^2 < 1. \quad (6.36)$$

This inequality can be rewritten in terms of one dimensional projections using the Bloch sphere representation of qubit states,

$$|\Phi_i(1)\rangle\langle\Phi_i(1)| = \frac{1}{2}\left(\mathbb{1} + \vec{x}_i \cdot \vec{\sigma}\right),$$

where $\vec{x}_i$ is a unit vector on the Bloch sphere and $\vec{\sigma}$ is the vector of Pauli matrices. We then have

$$\begin{aligned} |\langle\Phi_i(1)|\Phi_j(1)\rangle|^2 &= \mathrm{tr}(|\Phi_i(1)\rangle\langle\Phi_i(1)||\Phi_j(1)\rangle\langle\Phi_j(1)|) \\ &= \frac{1}{4}\left(\mathrm{tr}\,\mathbb{1} + \mathrm{tr}(\vec{x}_i \cdot \vec{\sigma}\vec{x}_j \cdot \vec{\sigma})\right) = \frac{1}{2}\left(1 + \vec{x}_i \cdot \vec{x}_j\right). \end{aligned} \quad (6.37)$$

Here we used the fact that the Pauli matrices are traceless. With this, equation (6.36) reads

$$\vec{x}_1 \cdot \vec{x}_2 + \vec{x}_1 \cdot \vec{x}_3 + \vec{x}_2 \cdot \vec{x}_3 < -1. \quad (6.38)$$

The invariance of the Haar measure means that the $\vec{x}_i$, $i = 1, 2, 3$, are three uniformly distributed unit vectors on the Bloch sphere. The integration over all bases using the Haar measure such that the inequality (6.38) holds yields the probability to be outside that tetrahedron face,

$$\begin{aligned} P_{\text{out,face}} &= \int d\vec{x}_1 d\vec{x}_2 d\vec{x}_3 \chi(\vec{x}_1, \vec{x}_2, \vec{x}_3), \\ \chi(\vec{x}_1, \vec{x}_2, \vec{x}_3) &= \begin{cases} 1 & \text{if } \vec{x}_1 \cdot \vec{x}_2 + \vec{x}_1 \cdot \vec{x}_3 + \vec{x}_2 \cdot \vec{x}_3 < -1, \\ 0 & \text{otherwise.} \end{cases} \end{aligned} \quad (6.39)$$

For the integration we will use that the scalar products of uniformly distributed unit vectors with a fixed vector are uniformly distributed on the interval $[-1, 1]$.

6.3.1 Proposition. *Let $d\vec{x}$ denote the invariant probability measure on the unit sphere in $\mathbb{R}^3$. Then, for any scalar function f and any unit vector $\vec{e}$,*

$$\int d\vec{x}\, f(\vec{x} \cdot \vec{e}) = \frac{1}{2}\int_{-1}^{1} dt\, f(t).$$

Proof. Since the measure is invariant under rotations of $\vec{x}$, we can take $\vec{e} = (0, 0, 1)$. Using the polar coordinates as in Figure 6.5, we get

$$\int d\vec{x}\, f(\vec{x} \cdot \vec{e}) = \frac{1}{4\pi} \int_0^{\pi} \sin\theta\, d\theta \int_0^{2\pi} d\varphi\, f(\cos\varphi).$$

The substitution $t = \cos\theta$ leads us to

$$\int d\vec{x}\, f(\vec{x} \cdot \vec{e}) = -\frac{1}{2} \int_{\cos 0}^{\cos\pi} dt\, f(t) = \frac{1}{2} \int_{-1}^{1} dt\, f(t),$$

which completes the proof. ∎

If we integrate (6.39) with respect to $\vec{x}_3$, keeping the other vectors fixed, the condition of the indicator function χ can be written in the form

$$(\vec{x}_1 + \vec{x}_2) \cdot \vec{x}_3 < -(1 + \vec{x}_1 \cdot \vec{x}_2).$$

Thus, we can apply Proposition 6.3.1 with $\vec{e} = (\vec{x}_1 + \vec{x}_2)/|\vec{x}_1 + \vec{x}_2|$ to equation (6.39),

$$P_{\text{out,face}} = \frac{1}{2} \int d\vec{x}_1 d\vec{x}_2 \int_{-1}^{1} dt f(\vec{x}_1, \vec{x}_2, t),$$
$$f(\vec{x}_1, \vec{x}_2, t) = \begin{cases} 1 & \text{if } t < -\frac{1+\vec{x}_1\cdot\vec{x}_2}{|\vec{x}_1+\vec{x}_2|}, \\ 0 & \text{otherwise.} \end{cases}$$

This amounts to

$$P_{\text{out,face}} = \frac{1}{2} \int d\vec{x}_1 d\vec{x}_2 \left(1 - \frac{1 + \vec{x}_1 \cdot \vec{x}_2}{|\vec{x}_1 + \vec{x}_2|}\right) = \frac{1}{2} \int d\vec{x}_1 d\vec{x}_2 \left(1 - \sqrt{\frac{1 + \vec{x}_1 \cdot \vec{x}_2}{2}}\right), \tag{6.40}$$

where we used the fact that $|\vec{x}_1 + \vec{x}_2| = \sqrt{2(1 + \vec{x}_1 \cdot \vec{x}_2)}$. Next, we apply Proposition 6.3.1 again to integrate (6.40) with respect to $\vec{x}_2$ and obtain

$$\begin{aligned} P_{\text{out,face}} &= \frac{1}{4} \int d\vec{x}_1 \int_{-1}^{1} dt \left(1 - \sqrt{\frac{1+t}{2}}\right) = \int d\vec{x}_1 \left(\frac{1}{2} - \frac{1}{\sqrt{32}} \int_0^2 ds \sqrt{s}\right) \\ &= \int d\vec{x}_1 \left(\frac{1}{2} - \frac{1}{3}\right) = \frac{1}{6}. \end{aligned} \tag{6.41}$$

In the last equation, we used the normalization of the Haar measure for the integration over $\vec{x}_1$. Thus, the probability to be outside the region delimited by the tetrahedron face $p_{12} + p_{13} + p_{23} \geq 1/2$ is $1/6$.

From the symmetry of the problem it is clear that this also holds for the other three faces of the tetrahedron that borders the region of triples that admit a classical model. Using the Bloch sphere representation (6.37), the inequalities for the

remaining three faces of the tetrahedron (6.31) can be written in the form

$$\begin{aligned} & s_{12}(1+\vec{x}_1\cdot\vec{x}_2)+s_{13}(1+\vec{x}_1\cdot\vec{x}_3)+s_{23}(1+\vec{x}_2\cdot\vec{x}_3)>2 \\ \Leftrightarrow\quad & s_{12}\vec{x}_1\cdot\vec{x}_2+s_{13}\vec{x}_1\cdot\vec{x}_3+s_{23}\vec{x}_2\cdot\vec{x}_3>2-(s_{12}+s_{13}+s_{23}), \end{aligned}$$

where s_{ij} is the sign of the term in the original inequality describing the tetrahedron face. As we always have the combination of one minus sign and two plus signs, the right hand side of the last inequality always amounts to 1. Furthermore, due to the invariance of the Haar measure, we can always change the sign of two terms on the left hand side, while keeping the sign of the third term. Thus we can change all signs on the left hand side to minus. By multiplication with (-1) we obtain (6.38), so this is exactly the same integration as for the first face.

This means that for random bases, drawn according to the Haar measure, the overall probability that the triples of marginals are not inside the classical tetrahedron (Figure 6.4) is $4/6 = 2/3$. So with probability $1/3$, a classical model exists, and therefore a retrodiction strategy for Alice.

However, her chance to survive is larger. Even if she cannot find a successful strategy in every case, she may be able to find the correct answer in some of the cases. This situation was analyzed using semidefinite programming (see subsection 6.3.4 below). That is, we have $\sum_x p(x)|\eta_x\rangle\langle\eta_x| < \mathbb{1}$ and her success probability is bounded below by $\sum_x p(x)$. For random bases, drawn according to the Haar measure, the average was $\langle\sum_x p(x)\rangle = 2/3$ for a sample size of ten million bases.

6.3.2 Expectation Maximization

One approach to find a successful retrodiction strategy for Alice is to consider the finding of a classical probability distribution as a hidden variable problem. Given the marginals (6.17), Gill [115] suggests to find a corresponding classical model (if it exists) via the expectation maximization (EM) algorithm [116]. To this end, we rewrite the marginals $p_{ab}(i,j)$ as probability distribution over the set

$$\{(a,b,i,j)\,|a<b; a,b\in\{1,\ldots,k\}; i,j\in\{1,\ldots,d\}\}$$

with probabilities

$$\tilde{p}(a,b,i,j) = Np_{ab}(i,j), \text{ where } N := \frac{1}{\sum_{a<b}1} = \frac{1}{\frac{k}{2}(k-1)}.$$

Clearly, $\tilde{p}(a,b,i,j)\geq 0$ and $\sum_{a<b}\sum_{i,j}\tilde{p}(a,b,i,j)=1$. The probabilities $p(l_1,\ldots,l_k)$ of the classical model now become hidden variables. The corresponding EM algorithm then reads:

Algorithm 1 Meaner-King-EM-Algorithm

Init: Start with a valid probability mass function, e. g., $q_1(l_1, \ldots, l_k) = \frac{1}{d^k}$, $l_i = 1, \ldots, d$, and compute the marginals

$$q_{ab}(i,j) = \sum_{l_1,\ldots,l_k=1}^{d} \delta_{l_a,i}\delta_{l_b,j} q(l_1, \ldots, l_k).$$

Step:

1. For each $a < b$ and each $(l_1, \ldots, l_k)$ define

$$q_{n+1}(l_1, \ldots, l_k) = N q_n(l_1, \ldots, l_k) \sum_{a<b} \frac{p_{ab}(l_a, l_b)}{q_{ab}(l_a, l_b)}. \tag{6.42}$$

2. Update the marginals $q_{ab}(i,j)$.

3. Compute the relative entropy between the marginals,

$$D(Np_{ab}(i,j) \| Nq_{ab}(i,j)) = N \sum_{a<b} \sum_{i,j} p_{ab}(i,j) \log\left(\frac{p_{ab}(i,j)}{q_{ab}(i,j)}\right).$$

Until: If the relative entropy is below a given threshold or if the gain to the previous relative entropy is above a given threshold, continue with the next step. Otherwise q is the best classical model found.

From equation (6.42) it follows immediately that a classical model q with marginals equal to $p_{ab}(i,j)$ is a fixed point of the iteration. Furthermore, convergence of the EM algorithm was shown in [116].

Looking at the iteration step, we see that the calculation of all $q_{ab}(i,j)$ as well as the calculation of all $q_{n+1}(l_1, \ldots, l_k)$ needs $d^k \frac{k}{2}(k-1)$ floating point operations each, if we calculate the $d^2 \frac{k}{2}(k-1)$ ratios p_{ab}/q_{ab} in advance. However, the calculation of the $d^2 \frac{k}{2}(k-1)$ logarithms in the relative entropy is expensive, so we may only check the terminating condition after several steps.

Implementation details: The code can be vectorized in several places. For example, the marginals of the bases can be computed as:

```
% compute given marginals from bases
pijab = zeros(d,d,N,N);
for a = 1:N
    for b=(a+1):N
        pijab(:,:,a,b) = 1/d*abs(bases(:,:,a)'*bases(:,:,b))^2;
    end
end
```

The marginals of the classical model can be computed via **sum** and **squeeze**:

```
margQijab = zeros(d,d,N,N);
%% compute marginals from q
for a = 1:N
    for b = (a+1):N
        margQAB = q;
        for iB = 1:N
            % sum over all but a and b
            if (iB ≠ a) && (iB ≠ b)
                margQAB = sum(margQAB, iB);
            end
        end
        margQijab(:,:,a,b) = squeeze(margQAB);
    end
end
```

Special attention has to be paid for the cases $q_{ab}(i,j) = 0$ and $0 \log 0 = 0$ in the iteration step. A linear index x is used for the probability distribution $q(l_1, \ldots, l_k)$, and the d-adic number representations of the indices $x = [l_1, \cdots, l_k]$ are precomputed.

Advantages of the algorithm are that it has a small memory footprint and is easy to implement. Disadvantages are that it converges slowly and that there is no way to penalize boundary solutions, if the solution set is open. Due to its poor speed of convergence this algorithm was dropped.

6.3.3 Linear Programming

Given the marginals (6.17), the existence of a classical model is a linear program feasibility problem. We are looking for a probability mass function $q(l)$, $l = (l_1, \ldots, l_k)$, with the marginals $p_\alpha = p_{ab}(i,j)$, $\alpha = (a,b,i,j)$. Thus, the feasibility problem can be written as

$$\begin{aligned} \text{minimize} \quad & 0 \\ \text{subject to} \quad & \sum_l q(l)\delta_\alpha(l) = p_\alpha \quad \forall\alpha \\ & q(l) \geq 0 \quad \forall l \\ & \sum_l q(l) = 1, \end{aligned} \tag{6.43}$$

where $\delta_\alpha(l) = \delta_{l_a,i}\delta_{l_b,j}$. Problem (6.43) is equivalent to the dual form,

$$\begin{aligned} \text{maximize} \quad & 0 \\ \text{subject to} \quad & A^*y + c = 0 \\ & y \geq 0, \end{aligned} \tag{6.44}$$

of the linear program

$$\begin{array}{ll} \text{minimize} & \langle c|x\rangle \\ \text{subject to} & Ax \leq 0. \end{array} \tag{6.45}$$

Applying Farka's Lemma (see [39]), we know that (6.44) is feasible if and only if the system of inequalities

$$Ax \leq 0, \quad \langle c|x\rangle < 0,$$

is infeasible.

To get a non-boundary solution in the case that the solution set is open, we consider the penalty function

$$f(q) := \frac{1}{2}\sum_l q(l)^2. \tag{6.46}$$

A more mixed probability distribution q has a smaller value of the penalty function. Hence our optimization problem becomes:

$$\begin{array}{lll} \text{minimize} & \frac{1}{2}\sum_l q(l)^2 & \\ \text{subject to} & \sum_l q(l)\delta_\alpha(l) = p_\alpha & \forall\alpha \\ & q(l) \geq 0 & \forall l \\ & \sum_l q(l) = 1. & \end{array} \tag{6.47}$$

6.3.4 Semidefinite Programming

Finding a classical model for the given bases can also be stated as semidefinite program. To this end we rewrite the condition (6.11) as matrix equation

$$M|\eta_x\rangle = |m_x\rangle, \tag{6.48}$$

where the matrix M and the vector $|m_x\rangle$ are defined as

$$M := \begin{pmatrix} \langle\Omega| \\ \langle\hat{\Phi}_1(1)| \\ \vdots \\ \langle\hat{\Phi}_1(d-1)| \\ \langle\hat{\Phi}_2(1)| \\ \vdots \end{pmatrix} \quad \text{and} \quad |m_x\rangle := \begin{pmatrix} 1 \\ \delta_{1,x(1)} \\ \vdots \\ \delta_{d-1,x(1)} \\ \delta_{1,x(2)} \\ \vdots \end{pmatrix}.$$

The first row of M and m_x describe our choice for the normalization, $\langle\Omega|\eta_x\rangle = 1$. This also fixes the scalar products with the last vector $\Phi_a(d)$ for all bases $a = 1, \ldots, k$ since $\sum_i |\hat{\Phi}_a(i)\rangle = |\Omega\rangle$. The matrix M is $(k(d-1)+1) \times d^2$ dimensional. In the

tomographically complete case, we have $k = (d+1)$ and $\hat{\mathcal{R}}_{\mathbb{C}}^{\perp} = \{0\}$. Therefore, M is invertible and the solution for $|\eta_x\rangle$ is given by

$$|\eta_x\rangle = M^{-1}|m_x\rangle.$$

The overall normalization condition (6.15) for a classical model p therefore becomes $\sum_x p(x)|M^{-1}m_x\rangle\langle M^{-1}m_x| = \mathbb{1}_{\hat{\mathcal{R}}}$, or equivalently

$$\sum_x p(x)|m_x\rangle\langle m_x| = M\mathbb{1}_{\hat{\mathcal{R}}}M^* = MM^*. \tag{6.49}$$

Note that the matrices $|m_x\rangle\langle m_x|$ are sparse. Since $\langle\Omega|\eta_x\rangle = 1$, we have

$$1 = \langle\Omega|\Omega\rangle = \langle\Omega|\sum_x p(x)|\eta_x\rangle\langle\eta_x||\Omega\rangle = \sum_x p(x)|\langle\Omega|\eta_x\rangle|^2 = \sum_x p(x)$$

by construction. Together with (6.14) we conclude that equation (6.49) implies that p is a probability mass function of a classical model if the $p(x)$ are positive for all x.

This means that finding a classical model is a semidefinite feasibility problem in the $p(x)$,

$$\begin{aligned} &\text{minimize} && 0 \\ &\text{subject to} && \sum_x p(x)|m_x\rangle\langle m_x| = MM^*, \\ & && p(x) \geq 0, \qquad x = 1, \ldots, d^k. \end{aligned} \tag{6.50}$$

This problem type,

$$\begin{aligned} &\text{minimize} && \langle c|z\rangle \\ &\text{subject to} && Az = b \\ & && z \geq 0, \end{aligned} \tag{6.51}$$

is also known as cone program in standard form [39] with the dual problem

$$\begin{aligned} &\text{maximize} && \langle b|y\rangle \\ &\text{subject to} && A^*y \leq c. \end{aligned} \tag{6.52}$$

Note that $\leq$ in (6.52) is the partial ordering of the dual cone, where $\geq$ in (6.51) is the partial ordering of the primal cone. If there is a solution $z > 0$ with $Az = b$, i. e., the Slater condition holds, we have strong duality. The alternative system is thus

$$\langle b|y\rangle > 0, \quad A^*y \leq 0. \tag{6.53}$$

In the case that one cannot find a classical model, i. e., a strategy for Alice that works for every choice of basis of the king, we are interested in to at least maximize Alice's survival chance. Given we found a solution with $\sum_x p(x)|\eta_x\rangle\langle\eta_x| < \mathbb{1}$, we can extend this to a POVM by the additional measurement operator $F_0 := \mathbb{1} - \sum_x p(x)|\eta_x\rangle\langle\eta_x|$.

Alice success probability P, given the king's choice a, is then bounded below by the sum of probabilities of the outcomes (i, x),

$$P := \sum_{x\neq 0}\sum_{i}\langle\hat{\Phi}_a(i)|p(x)|\eta_x\rangle\langle\eta_x||\hat{\Phi}_a(i)\rangle = \sum_{x\neq 0}\sum_{i}\delta_{i,x(a)}p(x) = \sum_{x\neq 0}p(x). \tag{6.54}$$

In fact, this bound on the success probability is independent of the king's choice a. The corresponding semidefinite problem is given by

$$\begin{aligned} &\text{maximize} && \sum_x p(x) && \\ &\text{subject to} && p_x \geq 0 && x = 1, \ldots, d^k \\ & && \sum_x p(x)|m_x\rangle\langle m_x| \leq MM^*. && \end{aligned} \tag{6.55}$$

6.3.4.1 Implementation Details

The optimization (6.55) is implemented in Matlab as conic program using SeDuMi [47] as solver. SeDuMi expects the input in standard conic primal form (6.51) or dual form (6.52). Therefore, the implementation of (6.55) in dual form is straight forward:

```
function y = SolveMeanKing( MMStar )
% Returns optimal probability mass function p(x) for the mean king
% problem.

% MMStar is (d^2)x(d^2)-matrix
d = sqrt(size(MMStar, 1));
N = d + 1;

% vectors | m_x >
mx = CreateMX(d, N);

% cone K, c - A'*y in K
%
% p_x ≥ 0
% M*M' - \sum_x p_x | m_x >< m_x | ≥ 0
K.l = d^N;
K.s = [d^2];
K.scomplex = [1];
K.xcomplex = [1:d^N];
K.ycomplex = [];

% b = <1...1|
b = ones(d^N, 1);

% c= [ 0, M*M'].'
c = zeros(d^N + d^4, 1);
```

```
c((d^N + 1):end,1) = vec(MMStar);

% At = [ 1, -|m_x >< m_x| ].'
At = sparse([], [], [], d^N + d^4, d^N);
At(1:d^N,:) = -speye(d^N);
for iX = 1:d^N
    At(d^N+1:end, iX) = vec(mx(:,iX)*(mx(:,iX)'));
end

% parameters
pars.fid = 0;
pars.eps = 1e-4;

% SeDuMi call
[x, y, info] = sedumi(At', b, c, K, pars);

if ((info.pinf == 1) || (info.dinf == 1))
    warning('Infeasible problem!');
end
```

6.3.4.2 Test Cases

1. For the first test case we choose the mutually unbiased bases, for which we already know that a strategy for Alice exists. For $d = 2$ mutually unbiased bases are given by the eigenvectors of the Pauli bases, i. e., the test case takes the bases

$$\left\{\left\{\frac{1}{\sqrt{2}}\begin{pmatrix}1\\1\end{pmatrix}, \frac{1}{\sqrt{2}}\begin{pmatrix}1\\-1\end{pmatrix}\right\}, \left\{\frac{1}{\sqrt{2}}\begin{pmatrix}1\\i\end{pmatrix}, \frac{1}{\sqrt{2}}\begin{pmatrix}1\\-i\end{pmatrix}\right\}, \left\{\begin{pmatrix}1\\0\end{pmatrix}, \begin{pmatrix}0\\1\end{pmatrix}\right\}\right\}.$$

2. The second test case is given by the bases (6.33). We expect the existence and non-existence of a classical model according to equation (6.35).

6.3.5 Numerical Results

The conic program (6.55) was applied to low dimensional cases. Unfortunately, the case $d = 6$ was not solvable on the available computer systems.

Figure 6.6 shows the results for the parametrized qubit bases (6.33). As one can see, the lower bound on the success probability decreases with the distance to the face of the Bell inequality tetrahedron. Note that in the case $\theta = \pi/2$, the bases are degenerate as they lie in the $x - y$-plane of the Bloch sphere (see Figure 6.5).

The conic program was also applied to $d + 1$ random bases. The randomness was generated by applying the Gram-Schmidt orthogonalization procedure to a matrix

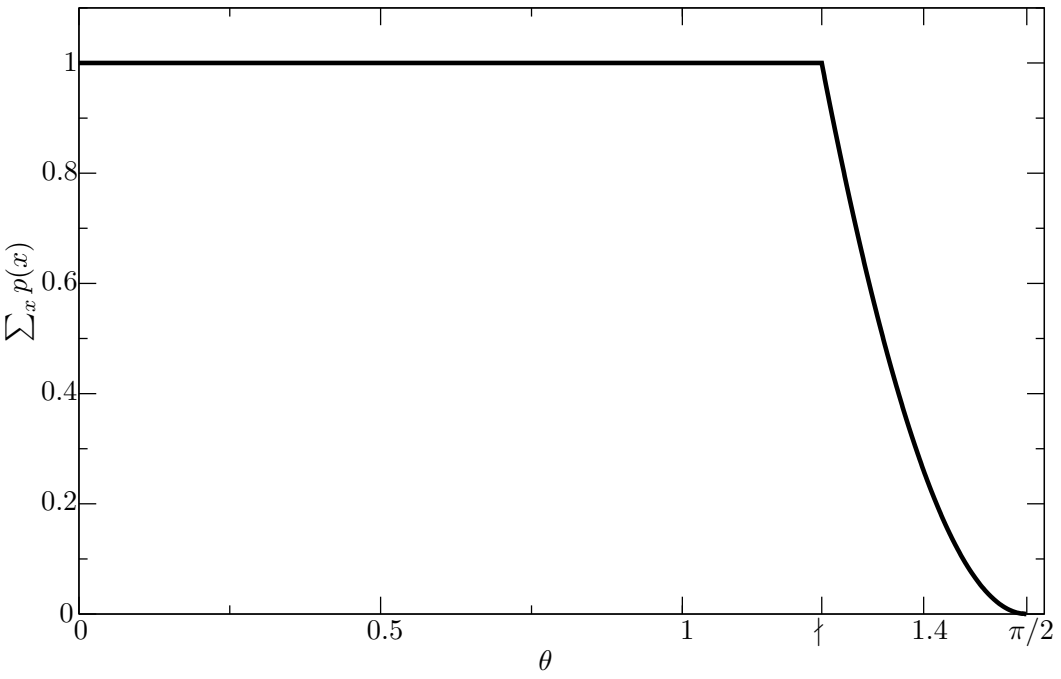

Figure 6.6: Lower bound on the probability to survive for Alice, if the king's bases are chosen according to (6.33). Here $\nmid$ marks the intersection of the p_{ab}-triples (6.34) with the tetrahedron due to the Bell inequality $p_{12} + p_{13} + p_{23} \geq 1/2$ of the tetrahedron (6.31). Inside the tetrahedron, a classical model and therefore a strategy for Alice exists for these bases, i. e.($\sum_x p(x) = 1$).

with pseudo random complex entries, where each entry consists of a real and imaginary part drawn from a normal distribution with mean 0 and standard deviation 1. The following table summarizes the numerical results.

d	P_s	$\langle \sum_x p(x) \rangle$	$\log_{10} N$
2	0.3342	0.6664	7
3	0.0013	0.398	6
4	0	0.35	4

Here P_s is the probability that a strategy exists, $\langle \sum_x p(x) \rangle$ is the average lower bound on the probability that she can find the right answer, and N is the sample size. In the case $d = 4$, none of the sample random bases admitted a classical model, although Alice overall success probability was about 0.35. These results suggest that the existence of a classical model in higher dimensions is rather exceptional.

6.4 Conclusion

In this chapter, it was shown that the mean king problem cannot be solved in all cases, if the king is not restricted to use mutually unbiased bases as in the original statement of the problem. This means that the king can choose bases, such that Alice cannot retrodict his measurement outcome. In the qubit case, a parametrized set of bases with this property was explicitly constructed and it was shown that for random bases, drawn according to the Haar measure, the probability for the existence of a strategy for Alice is 1/3. For any finite dimension, sufficient criteria for the existence of a strategy were given in the case of non-degenerate bases, and necessary conditions for the existence of a strategy were given in the case of tomographically complete bases. These criteria can be useful in the construction of security protocols, e. g., if the king must be able to deny his measurement result, but the chosen measurement apparatus becomes known by the attacker. Several numerical algorithms to decide whether a classical model exists for given marginals have been discussed. The numerical studies in higher dimensions suggest that only few cases admit a classical model, and therefore a successful retrodiction strategy for Alice. An interesting further study would be to decide whether the mean king protocol can be used to generate a secret key between the king and Alice, in the cases where a successful retrodiction strategy exists.

Appendix A

Tomography Listings

A.1 States

The following Matlab function computes the minimal least squares sum and a corresponding state for given tomography data. It is an implementation of the conic program (4.9) on page 105.

```
function [leastSquaresSum, fittedState] = StateEstimation(tomographyData)
%Least squares fit to tomograph results.
%
% tomographyData : Cell with tomography data in the form
%   {{A_1/sigma_1, a_1/sigma_1}, {A_2/sigma_2, a_2/sigma_2}, ... }
%
% return : Minimal least squares sum and fitted state. If the minimal
%   least squares sum is zero, then the tomography data already leads
%   to a positive semidefinite matrix with unit trace.

% get state dimension
% state is a dxd matrix
d = size(tomographyData{1}{1},1);

% number of terms in the least squares sum
n = size(tomographyData, 2);

% Dual problem in SeDuMi form:
% maximize Re b*y, subject to c - A*y is element of the cone

% create A and c
% (c - A*y) in cone
% At = A*

% Cone is R+ x R+ x Q x S x S x S
% tr(rho) - 1          in R+
```

```
% - tr(rho) + 1        in R+
% Least squares terms  in Q
% rho ≥ 0             in S
% i rho ≥ 0           in S
% -i rho ≥ 0          in S

% allocate memory
At = zeros(2 + 1 + n + 3*d^2, 1 + d^2);
c = zeros(2 + 1 + n + 3*d^2, 1);
b = zeros(1 + d^2, 1);

% == objective ==
b(1,1) = -1;

% == constraints ==
% === least squares ===
% y = (t, vec(rho))
% t is upper bound on least squares sum
% y is in the quadratic cone, t^2 ≥ least squares sum
At(3,1) = -1;

% === tomography constraints ===
%   tr ( rho A_i) - a_i) = (c - A*y)_j
%
%  trace(rho A) = sum_i sum_j <i| rho |j><j| A |i>
%
for ii=1:n
    Ai = tomographyData{ii}{1};
    ai = tomographyData{ii}{2};
    % At(3 + ii, 2:end) =
    for iii=1:d
        for jjj=1:d
            At(3 + ii, 1 + (jjj - 1)*d + iii) = Ai(jjj, iii);
        end
    end
    c(3 + ii, 1) = ai;
end

% === semidefinite state constraints ===
% tr (rho) = 1

% tr(rho) - 1 ≥ 0
% c - A*y in cone
c(1,1) = -1;
At(1,2:end) = -vec(eye(d))';

% -tr(rho) + 1 ≥ 0
% c - A*y in cone
c(2,1) = 1;
At(2,2:end) = vec(eye(d))';
```

```
% rho ≥ 0

% At( index(i,j) ) = MatrixOne(i,j)
%
% index(row i, column j) = (j − 1)*d + i
for ii = 1:d
    for jj = 1:d
        At((2+1+n) + (jj − 1)*d + ii, 1 + (jj − 1)*d + ii) = −1;
        % i*...
        At((2+1+n+d^2) + (jj − 1)*d + ii, 1 + (jj − 1)*d + ii) = −i;
        % −i*...
        At((2+1+n+2*d^2) + (jj − 1)*d + ii, 1 + (jj − 1)*d + ii) = i;
    end
end

% == solve ==
% specifiy cone
%
% note that the dual Scone is defined as
% { mat | mat + mat' ≥ 0 }
% hence we need
%    mat in dual Scone
%  i*mat in dual Scone
% −i*mat in dual Scone
% which results in mat ≥ 0

K.l = 2;
K.q = [1 + n];
K.s = [d, d, d];
K.xcomplex = [1:(K.q(1)+2)];
K.scomplex = [1, 2, 3];

% register complex variables
%
% bound and state diagonal terms are real
K.ycomplex = [];
for col = 1:(d−1)
    K.ycomplex = [K.ycomplex, ((col − 1)*d + col + 2):(col*d + col + 1) ];
end

% set misc SeDuMi options
pars.fid = 0;
pars.eps = 1e−6;

% SeDuMi call
[x, y, info] = sedumi(At', b, c, K, pars);

if ((info.pinf == 1) || (info.dinf == 1))
    warning('Infeasible problem!');
end
```

```
% set return value
leastSquaresSum = y(1);
fittedState = mat(y(2:end));
```

For a given upper bound t on the least squares sum, the Matlab function below computes the minimal and maximal pure state fidelity of a fitted state with a given pure state. This is an implementation of the conic program (4.10).

```
function [minFidelity, maxFidelity, minFittedState, maxFittedState] = ...
  MinMaxPureStateFidelity(pureState, leastSquaresSum, tomographyData)
% Optimizes the fidelity with a design pure state given a least squares
% distance to tomography data.
%
% The fidelity < pureState | fittedState | pureState >
% is optimized over all states with a given least squares sum.
%
% pureState : densitiy matrix of pure state as target for the objective
% leastSquaresSum : upper bound on the least squares sum, as, for example,
%   obtained from StateEstimation(tomographyData).
% tomographyData : Cell with tomography data in the form
%   {{A_1/sigma_1, a_1/sigma_1}, {A_2/sigma_2, a_2/sigma_2}, ... }
% return : maximum fidelity and fitted state that achieves it

% get state dimension
% state is a dxd matrix
d = size(tomographyData{1}{1},1);

% number of terms in the least squares sum
n = size(tomographyData, 2);

% Dual problem in SeDuMi form:
% maximize Re b*y, subject to c - A*y is element of the cone

% create A and c
% (c - A*y) \in cone
% At = A*

% Cone is R+ x R+ x Q x S x S x S
% tr(rho) - 1          in R+
% - tr(rho) + 1        in R+
% Least squares terms  in Q
% rho ≥ 0             in S
% i rho ≥ 0           in S
% -i rho ≥ 0         \in S

% allocate memory
At = zeros(2 + 1 + n + 3*d^2, d^2);
c = zeros(2 + 1 + n + 3*d^2, 1);
b = zeros(d^2, 1);
```

```
% == objective ==
% b'*y = <psi| rho |psi> = trace( rho |psi><psi| )
for ii=1:d
    for jj=1:d
        b((jj - 1)*d + ii) = pureState(jj, ii);
    end
end

% == constraints ==
% y = vec(rho)

% === tomography constraints ===
%  c - Ay = leastSquaresSum

c(3,1) = leastSquaresSum;

%   (tr(rho A_i) - a_i) = (c - A*y)_j
%
%  trace(rho A) = sum_i sum_j <i| rho |j><j| A |i>
%
for ii=1:n
    Ai = tomographyData{ii}{1};
    ai = tomographyData{ii}{2};
    for iii=1:d
        for jjj=1:d
            At(3 + ii, (jjj - 1)*d + iii) = Ai(jjj, iii);
        end
    end
    c(3 + ii, 1) = ai;
end

% === semidefinite state constraints ===
% tr(rho) = 1

% tr(rho) - 1 ≥ 0
% c - A*y in cone
c(1,1) = -1;
At(1,:) = -vec(eye(d))';

% -tr(rho) +1 ≥ 0
% c - A*y in cone
c(2,1) = 1;
At(2,:) = vec(eye(d))';

% rho ≥ 0

% At( index(i,j) ) = Matrixone(i,j)
%
% index(row i, column j) = (j - 1)*d + i
for ii = 1:d
    for jj = 1:d
```

```
        At((2+1+n) + (jj - 1)*d + ii, (jj - 1)*d + ii) = -1;
        % i*...
        At((2+1+n+d^2) + (jj - 1)*d + ii, (jj - 1)*d + ii) = -i;
        % -i*...
        At((2+1+n+2*d^2) + (jj - 1)*d + ii, (jj - 1)*d + ii) = i;
    end
end

% == solve ==
% specifiy cone
%
% note that the dual Scone is defined as
% { mat | mat + mat' ≥ 0 }
% hence we need
%    mat in dual Scone
%  i*mat in dual Scone
% -i*mat in dual Scone
% which results in mat ≥ 0

K.l = 2;
K.q = [1 + n];
K.s = [d, d, d];
K.xcomplex = [1:(K.q(1)+2)];
K.scomplex = [1, 2, 3];

% register complex variables
%
% bound and state diagonal terms are real
K.ycomplex = [];
for col = 1:(d-1)
    K.ycomplex = [K.ycomplex, ((col - 1)*d + col + 1):(col*d + col) ];
end

% set misc SeDuMi options
pars.fid = 0;
pars.eps = 1e-6;

% SeDuMi call
[x, y, info] = sedumi(At', b, c, K, pars);

if ((info.pinf == 1) || (info.dinf == 1))
    warning('Infeasible problem!')
end

% set return value
maxFidelity = real(b'*y);
maxFittedState = mat(y);

% == minimum fidelity ==
b = -b;
```

```
% SeDuMi call
[x, y, info] = sedumi(At', b, c, K, pars);

if ((info.pinf == 1) || (info.dinf == 1))
    warning('Infeasible problem!')
end

% set return value
minFidelity = -real(b'*y);
minFittedState = mat(y);
```

A.2 Channels

For given tomography data, the following Matlab function computes the minimal least squares sum, the minimal and maximal channel fidelity to a designated unitary gate, and the corresponding channels. It is an implementation of the conic program (4.12) with the additional channel fidelity optimizations.

```
function [ leastSquaresSum, minFidelity, maxFidelity, minChannel, ...
   maxChannel ] = UnitaryChannelFidelityLSFit( unitary, tomographyData )
% Returns minimum and maximum fidelity of least squares fit to channel
% tomography data compared to a unitary design.
%
% unitary : Design unitary. The fidelity that is optimized is
%   F_U(C) = trace( w (id tensor C)[(1 tensor U)w(1 tensor U)'])
%   where w is maximally entangled state, id the identity channel
%   and C a channel in Heisenberg picture.
% tomographyData : Tomography data in the form
%   {{rho_1, A_1/sigma_1, a_1/sigma_1}, {rho_2, A_2/sigma_2, a_2/sigma_2},
%     ... }.
% return : minimal least squares sum, minimum and maximum fidelity as well
%   as corresponding channels as HSChannel objects.

% == determine minimal least squares sum ==

% Channel
% C(a,i,b,j) == <a|C(|i><j|)|b>

% dimensions Channel: In -> Out
dIn = size(tomographyData{1}{2}, 1);
dOut = size(tomographyData{1}{1}, 1);

if dIn ~= dOut
    error('Input and output dimensions of the channel must be equal!');
end

% number of least squares terms
```

```
n = size(tomographyData, 2);

% matC = C( (a,i), (b,j) ) = reshape(C, dIn*dOut, dIn*dOut)
% reshape works columnwise:
% index(array(a,i,b,j)) = a
%                       + (i-1)*aMax
%                       + (b-1)*aMax*iMax
%                       + (j-1)*aMax*iMax*bMax

% Dual problem in SeDuMi form:
% maximize Re b*y
% subject to c - A'*y is element of the cone

% Scone = { mat | mat + mat' ≥ 0 }
% =>
% if
%      mat in S
%    i*mat in S
%   -i*mat in S
% then
%      mat ≥ 0

% y = (t, vec(matC))
% (t, least squares terms) in Q
% partialTrace(matC) - 1 ≥ 0
% -partialTrace(matC) + 1 ≥ 0
% matC ≥ 0
K.q = [1 + n];
K.s = [dOut, dOut, dOut, dOut, dIn*dOut, dIn*dOut, dIn*dOut];
K.scomplex = [1:7];
K.xcomplex = [1:K.q(1)];
K.ycomplex = [2:((dIn*dOut)^2 + 1)];

% allocate memory
At = zeros(1 + n + 4*dOut^2 + 3*(dIn*dOut)^2, 1 + (dIn*dOut)^2);
c = zeros(1 + n +  4*dOut^2 + 3*(dIn*dOut)^2, 1);
b = zeros(1 + (dIn*dOut)^2,1);

% === objective ===
b(1,1) = -1;

% === constraints ===

% ==== tomography data ====
At(1,1) = -1;

% tr( rho T(A) )
% = sum_b <b| rho T(A) |b>
% = sum_b sum_a <b| rho |a> <a| T(A) |b>
% = sum_b sum_a sum_i sum_j <b| rho |a> <a| T( |i><j|) |b> <i|A |j>
```

```
for iLSTerm = 1:n
    rhoi = tomographyData{iLSTerm}{1};
    Ai = tomographyData{iLSTerm}{2};
    % ai
    c(iLSTerm + 1, 1) = tomographyData{iLSTerm}{3};
    % trace(rhoi C(Ai)) - ai
    for iJ = 1:dIn
        for iB = 1:dOut
            for iI = 1:dIn
                for iA = 1:dOut
                    % reshape works columnwise:
                    % index(array(a,i,b,j)) = a
                    %                       + (i-1)*aMax
                    %                       + (b-1)*aMax*iMax
                    %                       + (j-1)*aMax*iMax*bMax
                    At(iLSTerm + 1, 1 + iA + (iI-1)*dOut + ...
                      (iB-1)*dOut*dIn + (iJ-1)*dOut*dIn*dOut) ...
                      = rhoi(iB, iA)*Ai(iI, iJ);
                end
            end
        end
    end
end

% ==== channel constraints ====
% c - At*y in K
%
% T(1) = 1
% sum_i sum_j Δ(i,j) <a| T( |i><j| ) |b> = Δ(a,b)

% paritalTrace(matC)  - 1 = 0
for iB = 1:dOut
    for iA = 1:dOut
        vtmp = zeros(dOut, dIn, dOut, dIn);
        vtmp(iA, :, iB, : ) = eye(dIn, dIn); % Δ(i,j)
        vtmp = reshape(vtmp, dOut*dIn, dOut*dIn);
        % paritalTrace(matC)  - 1 ≥ 0
        At(n + 1 + (iB - 1)*dOut + iA, :) = -[0; vec(vtmp)];
        % i*...
        At(n + 1 + dOut^2 + (iB - 1)*dOut + iA, :) = -i*[0; vec(vtmp)];
        % -i*...
        At(n + 1 + 2*dOut^2 + (iB - 1)*dOut + iA, :) = i*[0; vec(vtmp)];
        % -paritalTrace(matC)  + 1 ≥ 0
        At(n + 1 + 3*dOut^2 + (iB - 1)*dOut + iA, :) = [0; vec(vtmp)];
        % Δ(a,b)
        if (iA == iB)
            % paritalTrace(matC)  - 1 ≥ 0
            c(n + 1 + (iB - 1)*dOut + iA) = -1;
            % i*...
            c(n + 1 + dOut^2 + (iB - 1)*dOut + iA) = -i;
            % -i*...
```

```
            c(n + 1 + 2*dOut^2 + (iB - 1)*dOut + iA) = i;
            % -paritalTrace(matC)  + 1 >= 0
            c(n + 1 + 3*dOut^2 + (iB - 1)*dOut + iA) = 1;
        end
    end
end

% matC >= 0
for iJ = 1:dIn
    for iB = 1:dOut
        for iI = 1:dIn
            for iA = 1:dOut
                % reshape works columnwise:
                % index(array(a,i,b,j)) = a
                %                       + (i-1)*aMax
                %                       + (b-1)*aMax*iMax
                %                       + (j-1)*aMax*iMax*bMax
                At(1 + n + 4*dOut^2 + iA + (iI - 1)*dOut + ...
                  (iB - 1)*dOut*dIn + (iJ - 1)*dOut*dIn*dOut, ...
                  1 + iA + (iI - 1)*dOut + (iB - 1)*dOut*dIn + ...
                  (iJ - 1)*dOut*dIn*dOut) = -1;
                % i*...
                At(1 + n + 4*dOut^2 + dOut^2*dIn^2 + iA + ...
                  (iI - 1)*dOut + (iB - 1)*dOut*dIn + ...
                  (iJ - 1)*dOut*dIn*dOut, 1 + iA + (iI - 1)*dOut + ...
                  (iB - 1)*dOut*dIn + (iJ - 1)*dOut*dIn*dOut) = -i;
                % -i*...
                At(1 + n + 4*dOut^2 + 2*dOut^2*dIn^2 + iA + ...
                  (iI - 1)*dOut + (iB - 1)*dOut*dIn + ...
                  (iJ - 1)*dOut*dIn*dOut, 1 + iA + (iI - 1)*dOut + ...
                  (iB - 1)*dOut*dIn + (iJ - 1)*dOut*dIn*dOut) = i;
            end
        end
    end
end

% set misc SeDuMi options
pars.fid = 0;
pars.eps = 1e-10;

% SeDuMi call
[x, y, info] = sedumi(At', b, c, K, pars);

if ((info.pinf == 1) || (info.dinf == 1))
    warning('Infeasible problem!');
end

leastSquaresSum = y(1,1);

% == optimize fidelity with given unitary ==
% fix least squares sum to minimal value
```

```
c(1) = leastSquaresSum;
% remove upper bound coordinates
At = At(:,2:end);
K.ycomplex = [1:(dIn*dOut)^2];
% change objective
b = zeros((dIn*dOut)^2,1);
% C(a,i,b,j) == <a|C(|i><j|)|b>
% index(array(a,i,b,j)) = a
%                         + (i-1)*aMax
%                         + (b-1)*aMax*iMax
%                         + (j-1)*aMax*iMax*bMax

% F_U(C) = 1/d^2 sum_{a,b,i,j} <a|C(|i><j|)|b><i|U|a><b|U'|j>
d = dIn;
for iA = 1:d
    for iB = 1:d
        for iI = 1:d
            for iJ = 1:d
                b(iA + (iI-1)*d + (iB-1)*d^2 + (iJ-1)*d^3) = ...
                  unitary(iI,iA)'*unitary(iJ,iB)/d^2;
            end
        end
    end
end

% === maximum fidelity ===
% SeDuMi call
[x, y, info] = sedumi(At', b, c, K, pars);
if ((info.pinf == 1) || (info.dinf == 1))
    warning('Infeasible problem!');
end

maxFidelity = abs(b'*y);
maxChannel = CHSChannel(reshape( mat(y), dOut, dIn, dOut, dIn));

% === minimum fidelity ===
b = -b;

% SeDuMi call
[x, y, info] = sedumi(At', b, c, K, pars);
if ((info.pinf == 1) || (info.dinf == 1))
    warning('Infeasible problem!');
end

minFidelity = abs(b'*y);
minChannel = CHSChannel(reshape( mat(y), dOut, dIn, dOut, dIn));
```

Bibliography

[1] M. Reimpell and R. F. Werner. *Iterative Optimization of Quantum Error Correcting Codes.* Phys. Rev. Lett. **94** (8), 080501 (2005).

[2] M. Reimpell, R. F. Werner and K. Audenaert. *Comment on "Optimum Quantum Error Recovery using Semidefinite Programming".* arXiv.org (quant-ph/0606059) (June 2006).

[3] M. Reimpell and R. F. Werner. *Meaner King uses Biased Bases.* Phys. Rev. A **75** (6), 062334 (2007).

[4] O. Gühne, M. Reimpell and R. F. Werner. *Estimating Entanglement Measures in Experiments.* Phys. Rev. Lett. **98** (11), 110502 (2007).

[5] W. K. Wootters and W. H. Zurek. *A single quantum cannot be cloned.* Nature **299**, 802–803 (October 1982).

[6] D. Kretschmann, D. Schlingemann and R. F. Werner. *The Information-Disturbance Tradeoff and the Continuity of Stinespring's Representation.* arXiv.org (quant-ph/0605009) (April 2006).

[7] A. R. Calderbank and Peter W. Shor. *Good quantum error-correcting codes exist.* Phys. Rev. A **54** (2), 1098–1105 (Aug 1996).

[8] A. M. Steane. *Simple quantum error-correcting codes.* Phys. Rev. A **54** (6), 4741–4751 (Dec 1996).

[9] D. Gottesman. *Stabilizer codes and quantum error correction,.* PhD thesis California Institute of Technology Pasadena, California May 1997.

[10] Emanuel Knill and Raymond Laflamme. *Theory of quantum error-correcting codes.* Phys. Rev. A **55** (2), 900–911 (Feb 1997).

[11] H. Barnum, E. Knill and M. A. Nielsen. *On quantum fidelities and channel capacities.* IEEE Trans. Inf. Theory **46** (4), 1317–1329 (July 2000).

[12] D. G. Cory, M. D. Price, W. Maas, E. Knill, R. Laflamme, W. H. Zurek, T. F. Havel and S. S. Somaroo. *Experimental Quantum Error Correction.* Phys. Rev. Lett. **81** (10), 2152–2155 (Sep 1998).

[13] Nicolas Boulant, Lorenza Viola, Evan M. Fortunato and David G. Cory. *Experimental Implementation of a Concatenated Quantum Error-Correcting Code.* Physical Review Letters **94** (13), 130501 (2005).

[14] Debbie W. Leung, M. A. Nielsen, Isaac L. Chuang and Yoshihisa Yamamoto. *Approximate quantum error correction can lead to better codes.* Phys. Rev. A **56** (4), 2567–2573 (Oct 1997).

[15] K. Banaszek, G. M. D'Ariano, M. G. A. Paris and M. F. Sacchi. *Maximum-likelihood estimation of the density matrix.* Phys. Rev. A **61**, 010304(R) (1999).

[16] Massimiliano F. Sacchi. *Maximum-likelihood reconstruction of completely positive maps.* Phys. Rev. A **63** (5), 054104 (Apr 2001).

[17] M. G. A. Paris, G. M. D'Ariano and M. F. Sacchi. *Maximum-likelihood method in quantum estimation.* In *Bayesian inference and maximum entropy methods in science and engineering* vol. 568 of *AIP Conf. Proc.* page 456 2001.

[18] R. L. Kosut, I. Walmsley and H. Rabitz. *Optimal Experiment Design for Quantum State and Process Tomography and Hamiltonian Parameter Estimation.* arXiv.org (quant-ph/0411093) (2004).

[19] P. A. Schilpp, Editor. *Albert Einstein als Philosoph und Naturforscher.* W. Kohlhammer Verlag Stuttgart 1951.

[20] J. S. Bell. *On the Einstein-Podolsky-Roden paradox.* Physics **1**, 195–200 (1964).

[21] Alain Aspect, Philippe Grangier and Gérard Roger. *Experimental Tests of Realistic Local Theories via Bell's Theorem.* Phys. Rev. Lett. **47** (7), 460–463 (Aug 1981).

[22] Alain Aspect, Philippe Grangier and Gérard Roger. *Experimental Realization of Einstein-Podolsky-Rosen-Bohm Gedankenexperiment: A New Violation of Bell's Inequalities.* Phys. Rev. Lett. **49** (2), 91–94 (Jul 1982).

[23] Alain Aspect, Jean Dalibard and Gérard Roger. *Experimental Test of Bell's Inequalities Using Time- Varying Analyzers.* Phys. Rev. Lett. **49** (25), 1804–1807 (Dec 1982).

[24] M. Horodecki, P. Horodecki and R. Horodecki. *Separability of mixed states: necessary and sufficient conditions.* Phys. Lett. A **223** (1-2), 1–8 (November 1996).

[25] M.B. Plenio and S. Virmani. *An introduction to entanglement measures.* Quant. Inf. Comp. **7**, 1–51 (January 2007).

[26] Y. Aharonov and B.-G. Englert. *The mean king's problem: Spin 1.* Z. Naturforsch. **56a**, 16–19 (2001).

[27] B.-G. Englert and Y. Aharonov. *The mean king's problem: Prime degrees of freedom.* Phys. Lett. A **284** (1), 1–5 (2001).

[28] A. Klappenecker and M. Rötteler. *New Tales of the Mean King.* arXiv.org (quant-ph/0502138) (2005).

[29] G. Kimura, H. Tanaka and M. Ozawa. *Solution to the Mean King's problem with mutually unbiased bases for arbitrary levels.* Phys. Rev. A **73**, 050301(R) (2006).

[30] L. Vaidman, Y. Aharonov and D. Z. Albert. *How to ascertain the values of σ_x, σ_y, and σ_z of a spin-1/2 particle.* Phys. Rev. Lett. **58**, 1385 (1987).

[31] A. Hayashi, M. Horibe and T. Hashimoto. *Mean king's problem with mutually unbiased bases and orthogonal Latin squares.* Phys. Rev. A **71**, 052331 (2005).

[32] R. F. Werner. *Quantum Information* vol. 173 of *Springer Tracts in Modern Physics* chapter Quantum Information Theory - an Invitation, page 14. Springer Berlin May 2001.

[33] Lieven Vandenberghe and Stephen Boyd. *Semidefinite Programming.* SIAM Review **38** (1), 49–95 (1996).

[34] Michael Keyl. *Fundamentals of quantum information theory.* Physics Reports **369** (5), 431–548 (October 2002).

[35] Michael A. Nielsen and Isaac L. Chuang. *Quantum Computation and Quantum Information.* Cambridge Series on Information and the Natural Sciences. Cambridge University Press 2000.

[36] Ola Bratteli and Derek W. Robinson. *Operator Algebras and Quantum Statistical Mechanics I.* Texts and Monographs in Physics. Springer New York 1979.

[37] Ola Bratteli and Derek W. Robinson. *Operator Algebras and Quantum Statistical Mechanics II.* Texts and Monographs in Physics. Springer New York 1981.

[38] Vern I. Paulsen. *Completely Bounded Maps and Operator Algebras* vol. 78 of *Cambridge Studies in Advanced Mathematics.* Cambridge University Press Cambridge 2002.

[39] Stephen Boyd and Lieven Vandenberghe. *Convex Optimization*. Cambridge University Press 2004.

[40] Aharon Ben-Tal and Arkadi Nemirovski. *Lectures on Modern Convex Optimization* vol. 2 of *MPS/SIAM Series on Optimization*. Society for Industrial & Applied Mathematics Philadelphia 2001.

[41] R. Tyrrell Rockafellar. *Convex Analysis*. Princeton University Press Princeton, New Jersey 1970.

[42] J. Weidmann. *Lineare Operatoren in Hilberträumen*. B. G. Teubner Stuttgart 1976.

[43] W. Forrest Stinespring. *Positive Functions on C^*-Algebras*. Proc. Amer. Math. Soc. **6** (2), 211–216 (April 1955).

[44] Karl Kraus. *States, Effects, and Operations* vol. 190 of *Lecture Notes in Physics*. Springer Berlin 1983.

[45] Roger A. Horn and Charles R. Johnson. *Matrix Analysis*. Cambridge University Press Cambridge 1985.

[46] Michel X. Goemans and David Williamson. *Approximation algorithms for MAX-3-CUT and other problems via complex semidefinite programming*. In *STOC '01: Proceedings of the thirty-third annual ACM symposium on Theory of computing* page 443–452 New York 2001. ACM Press.

[47] J. F. Sturm. *Using SeDuMi 1.02, a MATLAB toolbox for optimization over symmetric cones*. Optimization Methods and Software **11-12**, 625–653 (1999).

[48] Peter W. Shor. *Polynomial-Time Algorithms for Prime Factorization and Discrete Logarithms on a Quantum Computer*. SIAM Journal on Computing **26** (5), 1484–1509 (1997).

[49] Peter W. Shor. *Scheme for reducing decoherence in quantum computer memory*. Phys. Rev. A **52** (4), R2493–R2496 (Oct 1995).

[50] A. M. Steane. *Error Correcting Codes in Quantum Theory*. Phys. Rev. Lett. **77** (5), 793–797 (Jul 1996).

[51] M. Keyl and R. F. Werner. *How to Correct Small Quantum Errors*. Lecture Notes in Physics **611**, 263–286 (2002).

[52] E. Knill. *Group Representations, Error Bases and Quantum Codes*. Technical Report LAUR-96-2807 Los Alamos National Laboratory 1996.

[53] M. Reimpell. *Fehlerkorrektur für depolarisierende Quantenkanäle.* Diplomarbeit 2002.

[54] D. Schlingemann. private communication 2007.

[55] Benjamin Schumacher. *Sending entanglement through noisy quantum channels.* Phys. Rev. A **54** (4), 2614–2628 (Oct 1996).

[56] Michał Horodecki, Paweł Horodecki and Ryszard Horodecki. *General teleportation channel, singlet fraction, and quasidistillation.* Phys. Rev. A **60** (3), 1888–1898 (Sep 1999).

[57] Raymond Laflamme, Cesar Miquel, Juan Pablo Paz and Wojciech Hubert Zurek. *Perfect Quantum Error Correcting Code.* Phys. Rev. Lett. **77** (1), 198–201 (Jul 1996).

[58] Charles H. Bennett, David P. DiVincenzo, John A. Smolin and William K. Wootters. *Mixed-state entanglement and quantum error correction.* Phys. Rev. A **54** (5), 3824–3851 (Nov 1996).

[59] D. Kretschmann and R. F. Werner. *Tema con variazioni: quantum channel capacity.* New J. Phys. **6**, 26 (February 2004).

[60] Claude Crépeau, Daniel Gottesman and Adam Smith. *Advances in Cryptology - EUROCRYPT 2005* vol. 3621 of *Lecture Notes in Computer Science* chapter Approximate Quantum Error-Correcting Codes and Secret Sharing Schemes, page 285. Springer 2005.

[61] H. Barnum and E. Knill. *Reversing quantum dynamics with near-optimal quantum and classical fidelity.* Journal of Mathematical Physics **43** (5), 2097–2106 (2002).

[62] Benjamin Schumacher and Michael D. Westmoreland. *Approximate quantum error correction.* arXiv.org (arXiv:quant-ph/0112106v1) (December 2001).

[63] K. Audenaert and B. De Moor. *Optimizing completely positive maps using semidefinite programming.* Phys. Rev. A **65**, 030302(R) (2002).

[64] A. Jamiolkowski. *Linear transformations which preserve trace and positive semidefiniteness of operators.* Rep. Math. Phys. **3** (4), 275–278 (December 1972).

[65] M.-D. Choi. *Completely positive linear maps on complex matrices.* Linear Algebra and its Applications **10**, 285–290 (1975).

[66] Beresford N. Parlett. *The symmetric eigenvalue problem.* Classics in Applied Mathematics 20. Prentice-Hall, Inc. Upper Saddle River, NJ, USA 1998.

[67] Gene H. Golub and Charles F. Van Loan. *Matrix computations (3rd ed.)*. Johns Hopkins University Press Baltimore, MD, USA 1996.

[68] J. H. Wilkinson and C. Reinsch. *Linear Algebra* vol. 186 of *Die Grundlehren der mathematischen Wissenschaften*. Springer Berlin 1971.

[69] Andrew S. Fletcher, Peter W. Shor and Moe Z. Win. *Optimum quantum error recovery using semidefinite programming*. Physical Review A (Atomic, Molecular, and Optical Physics) **75** (1), 012338 (2007).

[70] Robert L. Kosut and Daniel A. Lidar. *Quantum Error Correction via Convex Optimization*. arXiv.org (quant-ph/0606078v1) (June 2006).

[71] Emanuel Knill, Raymond Laflamme and Lorenza Viola. *Theory of Quantum Error Correction for General Noise*. Phys. Rev. Lett. **84** (11), 2525–2528 (Mar 2000).

[72] Paolo Zanardi. *Stabilizing quantum information*. Phys. Rev. A **63** (1), 012301 (Dec 2000).

[73] John A. Holbrook, David W. Kribs and Raymond Laflamme. *Noiseless Subsystems and the Structure of the Commutant in Quantum Error Correction*. Quantum Information Processing **2** (5), 381–419 (October 2003).

[74] Michael A. Nielsen. *A simple formula for the average gate fidelity of a quantum dynamical operation*. Physics Letters A **303** (4), 249–252 (October 2002).

[75] V. Bužek, M. Hillery and R. F. Werner. *Optimal manipulations with qubits: Universal-NOT gate*. Phys. Rev. A **60** (4), R2626–R2629 (Oct 1999).

[76] Michael Horodecki, Peter W. Shor and Mary Beth Ruskai. *Entanglement Breaking Channels*. Reviews in Mathematical Physics **15** (6), 629–641 (August 2003).

[77] H. Tverberg. *A generalization of Radon's theorem*. J. London Math. Soc. **41**, 123–128 (1966).

[78] M. Gregoratti and R. F. Werner. *On quantum error-correction by classical feedback in discrete time*. Journal of Mathematical Physics **45** (7), 2600–2612 (2004).

[79] N. Kiesel, C. Schmid, G. Toth, E. Solano and H. Weinfurter. *Experimental Observation of Four-Photon Entangled Dicke State with High Fidelity*. Physical Review Letters **98** (6), 063604 (2007).

[80] Heinz Bauer. *Wahrscheinlichkeitstheorie und Grundzüge der Maßtheorie*. Walter de Gruyter Berlin 1974.

[81] Daniel F. V. James, Paul G. Kwiat, William J. Munro and Andrew G. White. *Measurement of qubits.* Phys. Rev. A **64** (5), 052312 (Oct 2001).

[82] Mohamed Bourennane, Manfred Eibl, Christian Kurtsiefer, Sascha Gaertner, Harald Weinfurter, Otfried Gühne, Philipp Hyllus, Dagmar Bruß, Maciej Lewenstein and Anna Sanpera. *Experimental Detection of Multipartite Entanglement using Witness Operators.* Physical Review Letters **92** (8), 087902 (2004).

[83] J Eisert, F G S L Brandão and K M R Audenaert. *Quantitative entanglement witnesses.* New Journal of Physics **9** (3), 46 (2007).

[84] Koenraad M.R. Audenaert and Samuel L. Braunstein. *On Strong Superadditivity of the Entanglement of Formation.* Communications in Mathematical Physics **246** (3), 443–452 (April 2004).

[85] Ryszard Horodecki, Michał Horodecki and Paweł Horodecki. *Entanglement processing and statistical inference: The Jaynes principle can produce fake entanglement.* Phys. Rev. A **59** (3), 1799–1803 (Mar 1999).

[86] Fernando G. S. L. Brandao. *Quantifying entanglement with witness operators.* Physical Review A (Atomic, Molecular, and Optical Physics) **72** (2), 022310 (2005).

[87] Frank Verstraete and Michael M. Wolf. *Entanglement versus Bell Violations and Their Behavior under Local Filtering Operations.* Phys. Rev. Lett. **89** (17), 170401 (Oct 2002).

[88] S. P. Walborn, P. H. Souto Ribeiro, L. Davidovich, F. Mintert and A. Buchleitner. *Experimental determination of entanglement with a single measurement.* Nature **440** (7087), 1022–1024 (April 2006).

[89] Florian Mintert and Andreas Buchleitner. *Observable Entanglement Measure for Mixed Quantum States.* Physical Review Letters **98** (14), 140505 (2007).

[90] S.J. van Enk. *Can measuring entanglement be easy?* arXiv.org (quant-ph/0606017v1) (June 2006).

[91] Reinhard F. Werner. *Quantum states with Einstein-Podolsky-Rosen correlations admitting a hidden-variable model.* Phys. Rev. A **40** (8), 4277–4281 (Oct 1989).

[92] M. Barbieri, F. De Martini, G. Di Nepi, P. Mataloni, G. M. D'Ariano and C. Macchiavello. *Detection of Entanglement with Polarized Photons: Experimental Realization of an Entanglement Witness.* Phys. Rev. Lett. **91** (22), 227901 (Nov 2003).

[93] J. B. Altepeter, E. R. Jeffrey, P. G. Kwiat, S. Tanzilli, N. Gisin and A. Acin. *Experimental Methods for Detecting Entanglement.* Physical Review Letters **95** (3), 033601 (2005).

[94] Nikolai Kiesel, Christian Schmid, Ulrich Weber, Geza Toth, Otfried Gühne, Rupert Ursin and Harald Weinfurter. *Experimental Analysis of a Four-Qubit Photon Cluster State.* Physical Review Letters **95** (21), 210502 (2005).

[95] D. Leibfried, E. Knill, S. Seidelin, J. Britton, R. B. Blakestad, J. Chiaverini, D. B. Hume, W. M. Itano, J. D. Jost, C. Langer, R. Ozeri, R. Reichle and D. J. Wineland. *Creation of a six-atom 'Schrödinger cat' state.* Nature **438** (7068), 639–642 (December 2005).

[96] H. Häffner, W. Hänsel, C. F. Roos, J. Benhelm, D. Chek-al kar, M. Chwalla, T. Körber, U. D. Rapol, M. Riebe, P. O. Schmidt, C. Becher, O. Gühne, W. Dür and R. Blatt. *Scalable multiparticle entanglement of trapped ions.* Nature **438** (7068), 643–646 (December 2005).

[97] M. Lewenstein, B. Kraus, J. I. Cirac and P. Horodecki. *Optimization of entanglement witnesses.* Phys. Rev. A **62** (5), 052310 (Oct 2000).

[98] Dagmar Bruß, J. Ignacio Cirac, Pawel Horodecki, Florian Hulpke, Barbara Kraus, Maciej Lewenstein and Anna Sanpera. *Reflections upon separability and distillability.* Journal of Modern Optics **49** (8), 1399 – 1418 (July 2002).

[99] Geza Toth and Otfried Gühne. *Detecting Genuine Multipartite Entanglement with Two Local Measurements.* Physical Review Letters **94** (6), 060501 (2005).

[100] Otfried Gühne and Norbert Lütkenhaus. *Nonlinear Entanglement Witnesses.* Physical Review Letters **96** (17), 170502 (2006).

[101] Tzu-Chieh Wei and Paul M. Goldbart. *Geometric measure of entanglement and applications to bipartite and multipartite quantum states.* Phys. Rev. A **68** (4), 042307 (Oct 2003).

[102] R. F. Werner. *Statistische Thermodynamik.* Skript zur Vorlesung 1991.

[103] Abner Shimony. *Degree of Entanglement.* Annals of the New York Academy of Sciences **755** (1), 675–679 (1995).

[104] H Barnum and N Linden. *Monotones and invariants for multi-particle quantum states.* Journal of Physics A: Mathematical and General **34** (35), 6787–6805 (2001).

[105] T.-C. Wei, M. Ericsson, P. M. Goldbart and W. J. Munro. *Connections between relative entropy of entanglement and geometric measure of entanglement.* Quantum Information & Computation **4** (4), 252–272 (July 2004).

[106] M. Hayashi, D. Markham, M. Murao, M. Owari and S. Virmani. *Bounds on Multipartite Entangled Orthogonal State Discrimination Using Local Operations and Classical Communication.* Physical Review Letters **96** (4), 040501 (2006).

[107] W. K. Wootters and B. D. Fields. *Optimal State Determination By Mutually Unbiased Measurements.* Ann. Phys. (NY) **191** (2), 363–381 (1989).

[108] *Some Open Problems in Quantum Information Theory.* arXiv.org (quant-ph/0504166) (April 2005).

[109] P. K. Aravind. *Solution to the King's Problem in prime power dimensions.* Z. Naturforsch. **58a**, 2212 (2003).

[110] P. K. Aravind. *Best conventional solutions to the King's Problem.* Z. Naturforsch. **58a**, 682–690 (2003).

[111] O. Schulz, R. Steinhübl, M. Weber, B.-G. Englert, C. Kurtsiefer and H. Weinfurter. *Ascertaining the Values of σ_x, σ_y, and σ_z of a Polarization Qubit.* Phys. Rev. Lett. **90**, 177901 (2003).

[112] M. Horibe, A. Hayashi and T. Hashimoto. *Solution to the king's problem with observables that are not mutually complementary.* Phys. Rev. A **71**, 032337 (2005).

[113] S. Ben-Menahem. *Spin-measurement retrodiction.* Phys. Rev. A **39** (4), 1621–1627 (February 1989).

[114] Y Aharonov and L Vaidman. *Complete description of a quantum system at a given time.* Journal of Physics A: Mathematical and General **24** (10), 2315–2328 (1991).

[115] R. Gill. private communication 2006.

[116] A. P. Dempster, N. M. Laird and D. B. Rubin. *Maximum Likelihood from Incomplete Data via the EM Algorithm.* J. Roy. Stat. Soc. B **39**, 1–38 (1977).

Acknowledgments

I gratefully acknowledge the support from many people. First of all, I like to thank Prof. Dr. Reinhard F. Werner for mentoring this thesis, for introducing me to quantum information, for his tremendous support and the countless whiteboard discussions. I very much appreciate his sharing of experience and knowledge about mathematical physics, and it was a great pleasure to be the scientific assistant for his lectures.

I also like to thank PD Dr. Michael Keyl, who kindly agreed to review this thesis, for the discussions about mathematical physics and open source during his time in Braunschweig, and his helpful guide "Fundamentals of Quantum Information Theory" [34].

Thanks go to Dr. Koenraad M. R. Audenaert for the joint work and the discussions about semidefinite programming, Dr. Otfried Gühne for the joint work and the discussions about entanglement witnesses, and Jun.-Prof. Dr. Jens Eisert for the discussion about semidefinite relaxations and especially for his efforts trying to offer me a research position at Imperial College.

I very much enjoyed my stay at Prof. Dr. Harald Weinfurter's experimental quantum physics group at LMU and MPQ München in February 2006. Many thanks for the invitation and, in particular, to Nikolai Kiesel, Christian Schmid, and Witlef Wieczorek for the demonstration of optical experiments and the discussions about quantum tomography.

Nikolai Kiesel and Dr. Mark Riebe provided me with experimental test data for the tomography of quantum states and unitary gates. Thanks a lot!

Many thanks to the IMaPh group, to my part-time roommate PD Dr. Dirk Schlingemann for his algebraic view on quantum error correction and the joyful competition about the worst surprise in the corresponding eggs, to Ole Krüger for the discussions not only about physics and his encouraging words, to Dennis Kretschmann for his feedback and suggestions, to Torsten Franz for the non-smoker cigarette breaks and his comments about this thesis as well as about movies, to Annette Gattner and Holger Vogts for sharing the couch in their room with me and for their feedback on the thesis, to Albert Werner for his interest in the mean king problem and the

proof-reading, and last but not least to the rest of the group for the discussions during lunch. Many thanks to Cornelia “Conny” Schmidt, who masterly took care about all the administrative tasks.

Special thanks go to my wife Maren and my children Lara and Juliane for their love and patience. Maren gave me unstinting support throughout the time. Lara and Juliane showed me how much fun classical mechanics can be, during our joint experiments with building bricks.

Funding from the Deutsche Forschungsgemeinschaft (WE 1240/9), the TU Braunschweig, and the European Commission (FP6 STREP QICS) is gratefully acknowledged.